

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2002, 17:45:00 ; Search time 16860.4 seconds
 (without alignments)
 12121.202 Million cell updates/sec

Title: us-09-700-906a-1_copy_197_9962
 Perfect score: 9766
 Sequence: 1 atdttgcccacagacgcct.....tcatagggacagtgaagata 9766

Scoring table: OLIGO_NUC
 Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description

1	9766	100.0	12493	6	AX009576	
2	9715	99.5	12515	6	AX336950	
3	9715	99.5	12515	9	HSMKI67	
4	8286	84.8	11435	9	HSMKI67A	
5	6847	70.1	12965	9	HSAMAK67	
c 6	5227	53.5	123057	9	AL390236	
c 7	2855	29.2	140466	9	AL355529	
8	629	6.4	733	6	AX147668	
9	456	4.7	568	6	AX341136	
c 10	245	2.5	318	11	G33252	
c 11	199	2.0	332	11	G03247	
c 12	198	2.0	409	6	AX336882	
c 13	110	1.1	158	6	AX072812	
14	93	1.0	14041	9	HSKI67	
c 15	64	0.7	176433	9	HS71L16	
16	55	0.6	391	9	MMKI67391	
17	51	0.5	51	6	AX204486	
18	40	0.4	430	9	MMKI67430	
19	28	0.3	9471	10	MMKI67	
20	24	0.2	35540	8	SPBC36	
c 21	24	0.2	63170	10	AL645799	
c 22	24	0.2	161997	2	AL645526	
23	24	0.2	194575	2	AC023140	
24	24	0.2	229232	2	AC096624	
25	24	0.2	247899	2	AC025909	
c 26	23	0.2	23	6	AX009578	
27	23	0.2	184	4	BOVMICREB	
28	23	0.2	53419	10	MMU35323	
c 29	23	0.2	62280	2	AC094189	
30	23	0.2	71299	2	AC107772	
c 31	23	0.2	73198	9	HSDJ127L3	
32	23	0.2	73402	2	AC101535	
33	23	0.2	73886	2	AC106156	
c 34	23	0.2	78693	9	HSDJ475B7	
35	23	0.2	96951	9	AC026433	
36	23	0.2	100653	2	AC093215	
c 37	23	0.2	124814	9	AL161436	
c 38	23	0.2	125630	2	AC097551	
39	23	0.2	145157	2	AC103108	
40	23	0.2	148000	2	AC108922	
41	23	0.2	155397	9	AL590523	
42	23	0.2	156662	2	AC024454	
43	23	0.2	157639	2	AC103026	
44	23	0.2	164562	2	AC027471	
45	23	0.2	165700	2	AC099450	

ALIGNMENTS

RESULT	1					
AX009576						
LOCUS	AX009576	Sequence 1	from Patent	12493 bp	DNA	linear PAT 06-SEP-2000
DEFINITION	Sequence 1	from Patent	WO9961607.			
ACCESSION	AX009576					
VERSION	AX009576.1	GI:9996814				
KEYWORDS	.					
SOURCE	human.					
ORGANISM	Homo sapiens					

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 12493)
 Deinert,I., Boehle,A., Gerdes,J. and Flad,H.D.
 Antisense oligonucleotides for treating proliferating cells
 Patent: WO 9961607-A 1 02-DEC-1999;
 TITLE
 JOURNAL
 DEINERT IRINA (DE); BOEHLE ANDREAS (DE); GERDES JOHANNES (DE); FLAD
 HANS DIETER (DE); FORSCHUNGSZENTRUM BORSTEL ZENT (DE)

FEATURES
 Location/Qualifiers
 1..12493
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 197..9967
 /note="unnamed protein product"

Db 1337 GAAGGCTTCAAGGCTGGTGATAAAACATCTTACTCCAGGAAGCTTCAACTAGAAATCGA 1396
Qy 1201 acaccagctaaagtgaagatgcaagtgcagctgactctgccactaagccagagaaatctctcttc 1260
Db 1397 ACACCAGCTAAAGTTGAAGATGCAGCTGACTTGCCTACTAAGCCAGAAAATCTCTCTTCC 1456
Qy 1261 aaaccagaggaagtattctctacagatgtgtgaagttctgctactcggaactgaaattcac 1320
Db 1457 AAAACCCAGAGGAAGTATTCCTACAGATGTGGAAGTTCTGCCTACGGAAACTGAAATTCAC 1516
Qy 1321 aatgagccatttttaactctgtgctcactcaactgaattgagaggaagatccaaaagattcc 1380
Db 1517 AATGAGCCATTTTAACTCTGTGCTCCTCACTCAAGTTGAGAGGAAGATCCAAAAGGATTC 1576
Qy 1381 ctacgcaagcctgagaaattggcactcacagctgagacagatgtgctctgggttacctgggt 1440
Db 1577 CTCAGCAAGCCTGAGAAATTTGGGCACATACAGCTGACAGATGTCTCTGGGTTACCTTGGT 1636
Qy 1441 cttagttcagttgatatacaaaactttggtgattccattaatgagagtgagggaatacct 1500
Db 1637 CTTAGTTTCAGTTGATATCAACACTTTGGTGATTTCCATTAAATGAGAGTGAGGGAATACCT 1696
Qy 1501 ttgaaaagaaggcgtgtgctcttgggtgggcacacctaaagacctgaactatttggatgaaac 1560
Db 1697 TTGAAAAGAAGGCGTGTCTCTTTGGTGGGCACCTAAGACCTGAACATATTTGATGAAAC 1756
Qy 1561 ttgctcctaatagcctctcaaaaggggagaaagcccaacaaagaaagttctctggta 1620
Db 1757 TTGGCTCTTAATAGCCCTCTCAAAAGGGGAGAGCCCCAACCAAAAAGAAAGTCTCTGGTA 1816
Qy 1621 atgcacactccactgtctctgaagaaatcatcaaggaacagcctcaacatcaggaaaa 1680
Db 1817 ATGCACACTCCACCCTGCTCTGAAGAAATCATCAAGGAACAGCCCTCAACCATCAGGAAA 1876
Qy 1681 caagatcaggttcagaaatccatgtggaagtgaagcagacaaaacttgggttataagcct 1740
Db 1877 CAAGATCAGGTTCAAGAAATCCATGTGGAAGTGAAGGCACAAAGCTTGGTTTAAAGCCCT 1936
Qy 1741 ccagctcctagtctagaaaaactccagttgcccagtgatcaacgccgtagctctgcaaa 1800
Db 1937 CCAGCTCTAGTCTAGTCAAAACTCCAGTTGCCAGTGATCAACGCCGTAGTCTCGCAA 1996
Qy 1801 acagccctgcttccagcagcaaatctcagacagaggttctctaagagagagaggaaga 1860
Db 1997 ACAGCCCTGTCTCCAGCAGCAAAATCTCAGACAGAGGTTCTTAAGAGAGGAGGAGAAAG 2056
Qy 1861 gtgcgaacctgcttcaaaagagagtgctctatcagccgaagtcaacatgatattttacag 1920
Db 2057 GTGGCAACCTGCCTTCAAAAGAGAGTGTCTATCAGCCGAAGTCAACATGATATTTTACAG 2116
Qy 1921 atgatatgttccaaaagaagaagtgggtgcttcggaagcaaatctgattgttgcaaaatca 1980
Db 2117 ATGATATGTTCCAAAAGAGAAGTGGTGTCTCGGAAGCAAAATCTGATTTGTGCAAAATCA 2176
Qy 1981 tgggcagatgtagtataaacttgytgcaaaacaaacacaaactaaagtcatataaaca tgggt 2040
Db 2177 TGGGCAGATGTAGTAAAAACTTGGTGCAAAACAAAACATAAAGTCAATAAAACATGGT 2236
Qy 2041 cctcaaaaggtcaatgaaacaaaggcaagaagacctgctactctcaagaagcctgtgggc 2100
Db 2237 CCTCAAAAGGTCAATGAACAAAGGCAAGAAAGACCTGTACTCTCAAAAGAGCCGTGGGC 2296
Qy 2101 gaattcacagttcaatttagtacaggccacgcaaaactctcttgaccataataataggg 2160
Db 2297 GAAGTTCACAGTCAATTTAGTACAGGCCACGCAAACTCTCCTTGTACCATTAATAATAGGG 2356
Qy 2161 aaagctcactactgaaaaagtacatgtgctctgctcgacccctacagagtgctcaacaatttc 2220
Db 2357 AAAGCTCATACTGAAAAGTACATGTGCCCTGCTCGACCCCTACAGAGTGTCTCAACAACCTC 2416
Qy 2221 attttcaaccacaaaaatggagcttttaaggaagatctttcaggaatagctgaaattcaag 2280
Db 2417 ATTTCCAAACCAAAAAATGGAGCTTTTAAAGGAAGATCTTTTCAGGAATAGCTGAAATGTTCAAG 2476

Qy 2281 accccagtgaaagagcaaccgcagttgacaagcacatgtcacatcgctattttcaaatca 2340
Db 2477 ACCCCAGTGAAGGAGCAACCCGAGTTGACAGACACATGTACATCGCTATTTCAAAATCA 2536
Qy 2341 gagaatttgcctggaaaaacagtttcaaggaaactgattcaggagagaagaacctctgtctccc 2400
Db 2537 GAGAAATTTGCTTGGAAAACAGTTTCAAGGAATGATTTCTCAGGAGAAGAACCTCTGTCTCCC 2596
Qy 2401 acctcaagaagtttgaggaaatgtgttcttcagtgacagaatgcagcaaaaacagcca 2460
Db 2597 ACCTCAGAGAGTTTGGAGGAAATGTGTCTTCTAGTCACAGAAATGACAGCAAAACAGCCA 2656
Qy 2461 tctgataaattcctctgcaagccctcccttaagaagcgagtgatttagagaaaaatggaac 2520
Db 2657 TCTGATAAATCCTCTGCNAAGCCCTCCCTTAAGACGGCAGTGTATTAGAGAAAATGGAAC 2716
Qy 2521 gtagcaaaaaacgcccaggaacacactacaaaatgacttctctgagagacaaaacttcagat 2580
Db 2717 GTAGCAAAAACGCCAGGAACACCTTACAAAATGACTTCTCTGGAGACAAAACCTTCAGAT 2776
Qy 2581 actgagacagagccttcaaaaacagtatccactgtataacaggtcaggaaggtctacaag 2640
Db 2777 ACTGAGACAGAGCCCTTCAAAAACAGTATCCACTGTAAACAGGTCTAGGAAGGTCTACAGAG 2836
Qy 2641 ttcagggaatacacagaagctacctctgtgaaaagtgaagatgaagaacaaaacacagaaatt 2700
Db 2837 TTCAGGAATATACAGAAGCTACCTCTGGAAAGTAAAGAGTGAAGAAACAANAATACAGAAAT 2896
Qy 2701 gttgagtgcatacctataaaagggttcagaaggcaacactactacaacaaggagagaaga 2760
Db 2897 GTTGAGTGCATCTCTAAAAGAGGTGAGAAGGCACACACTACTACAAACAAGGAGAGAAGGA 2956
Qy 2761 gagatgaaggaaatagaaagaccttttgagacataaaggaaaattattgaattaaagaa 2820
Db 2957 GAGATGAAGGAAATAGAAAGACCTTTTGAGACATATAAGGAAAATATTGAAATTAAGAA 3016
Qy 2821 aacgatgaaaagatgaaagcaatgaagagatcaagaactctggggcagaaaattgtgcacca 2880
Db 3017 AACGATGAAAAGATGAAGCAATGAAGAGATCAAGAACTTGGGGCAGAAATGTGCACCA 3076
Qy 2881 atgtctgaactgcagacacctcaagagcttgctgatacagaactcatgaagaacacgcca 2940
Db 3077 ATGTCTGACCTTGACAGACCTCAAGAGCTTGGCTGTATACAGAACTCATGAAAGACACGCA 3136
Qy 2941 cgtggccagaatctcctccaaacccaagatcatgccaggcaccacaaagagtgagaagagc 3000
Db 3137 CGTGGCCAGAAATCTCCTCCAAACCCCAAGATCATGCCAAGGCCACCAAGAGTGAGAAAGGC 3196
Qy 3001 aaaatcactaaaaatgccctgccagtcattacaacccagacccaaataaacacccccaacacac 3060
Db 3197 AAATCACTAAATGCCCTGCCAGTCAATTACAACCCAGAACCAATAAACCCCAACACAC 3256
Qy 3061 acaaaacaaactgtgaaggcatccctggggaaagttagtgtgaaagaagagctcctagca 3120
Db 3257 ACAAAACAACAGTTGAAGGCATCCCTGGGAAAGTAGGTGTGAAAGAGAGGCTCTAGCA 3316
Qy 3121 gtgcgaagtccacagagcgtcagggagagccacacacacacacacagacagagccagcagga 3180
Db 3317 GTCGGCAAGTTTCACACGGACGTGAGGGGAGACCAGCACGCACAGAGAGCCAGCAGGA 3376
Qy 3181 gatggcaagagcatcagaacgctttaaggagtcctcaaaagcagatccttgaccaccagacc 3240
Db 3377 GATGCAAGACATCAGAACGTTTAAAGGAGTCTCCAAAGAGATCTCTGGACCCACGACC 3436
Qy 3241 cgttactactgaaatgaagaagtgcccaagaacgcctaaagggaagagcccgactcactagaa 3300
Db 3437 CGTGTAACTGGAATGAAGAGTGGCCAAAGACGCCTAAGGAAGAGGCCCGCTCACTAGAA 3496
Qy 3301 gacctggctgcttcaaaagactcttcagacacacagctccctctgaggaatcaatgact 3360
Db 3497 GACCTGGCTGCTTCAAAAGAGCTCTTCCAGACACACAGGTCCTCTGTAGGAATCAATGACT 3556

QY 3361 gatgagaaactaccaaaatagcctgcataatctccaccaccagaaatcagtggacactcca 3420
Db 3557 GATGAGAAAATACCAAAAATAGCTGCAAAATCTCCACCACCAGAAATCAGTGGACACTCCA 3616
QY 3421 acaagcacaaagcaatggcctaaagaagctctcagaaagcagatgtagaggaatctc 3480
Db 3617 ACNAGCACAAAGCAATGGGCTTAAGAAAGTCTCAGGAAAGCAGATGTAGAGGAAGAAATTC 3676
QY 3481 ttgactcaggaataaacaccatcagcagggaagccatgcttaccgcccacacagca 3540
Db 3677 TTAGCACTCAGGAAACTAACACCATCAGCAGGGAAGCCATGCTTACGCGCCCAACACGCA 3736
QY 3541 gagggtga tgaagaagacattaaagcatttatyggaaactccagtcagaaactggacctg 3600
Db 3737 GGAGGTGATGAGAAAGACATTTAAAGCATTTATGGGAATCCAGTGCAGAAACTGGACCTG 3796
QY 3601 cgaggaaacttaccctgcagcaaaaagacagctacagactcctaaggaaaggccagctc 3660
Db 3797 GCAGGAACCTTTACCTGGCAGCAAAAGACAGCTACAGACTCCTTAAGGAAAGGCCCAGGCT 3856
QY 3661 ctagaagacctggctggctttaagagctcttccagactcctgggtcacaccgagggaattc 3720
Db 3857 CTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACTCCTGGTCACACCGAGGAATTA 3916
QY 3721 gtggctgtctggtaaaaccactaaaaataccctgcgactctccacagtcagaccagtgagc 3780
Db 3917 GTGGCTGTCTGGTAAACCACTAAAATACCTCTCGCACTCTCCACAGTCAGACCCAGTGGAC 3976
QY 3781 acccaacaagcacaaagcaacgaccccaagagaagtatcaggaagcagatgtagaggga 3840
Db 3977 ACCCAACAAGCACAAAGCAAGCACCAGAGAAAGTATTAGGAAAGCAGATGTAGAGGGA 4036
QY 3841 gaactcttagctgtagggaatcttaatgccaatcagcaggcaaaagccatgcacacgcctaaa 3900
Db 4037 GAACTCTTAGCTGTCAGGAATCTAATGCCATCAGCAGGCAAGGCCATGCACACGCCCTAAA 4096
QY 3901 ccaatcagtaggtgaagagaaagacatcatcatattgtgtggaaactccagtcagaaactg 3960
Db 4097 CCATCAGTAGTGGTGAAGAGAAAGACATCATATTTGTGGGAATCCAGTGCAGAAACTG 4156
QY 3961 gacctgacagaaacttaaccgcgacgaagacagcgccacaactccttaagaaaggcc 4020
Db 4157 GACCTGACAGAGAACTTAACCGCGACGAAGAGACGGCCCAAACTCCTTAAGGAAGAGGCC 4216
QY 4021 caggctctggaagacactgactggctttaagagctcttccagaccctgggtcatactgaa 4080
Db 4217 CAGGCTCTGGAAGACCTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTCATCTGAA 4276
QY 4081 gaagcagtggtctgtggcaaaactactaaaaatgcccctgcgaatcttctccaccagaatca 4140
Db 4277 GAAGCAGTGGCTGCTGGGCAAAACTACTAAAATGCCCTGGCAATCTTCTCCACAGAAATCA 4336
QY 4141 gcagacaccccaacgaacgaagaaggcagcccaagacacctttggagaaaggagcgta 4200
Db 4337 GCAGACNCCCAACAAAGCACAGAGAGGCGAGGCCCAAGACACACTTTGGAGAAAAGGGACGTA 4396
QY 4201 cagaaggagctctcagccctgaagaagctcacacagacatcaggggaaaccacacacaca 4260
Db 4397 CAGAAGGAGCTCTCAGCCCTGAAGAAGCTCACACAGACATCAGGGGAAACCCACACACACA 4456
QY 4261 gataaagtaccaggaggtgagatataaagcatcaacgcgctttagggaaactcctcaaaacg 4320
Db 4457 GATAAAGTACCAGGAGGTGAGGATTAAGACATCAACGCGTTTATAGGAAACTGCAAAACAG 4516
QY 4321 aaactggaccagcagcaagttaactggtagcaagagccacccaaaactaaggaagaag 4380
Db 4517 AAATGGACCCAGACAAAGTGTAACTGGTAGCAAGAGGCACCCAAAAACTAAGGAAAAG 4576
QY 4381 gcccaacccctagaagacctggctgtggaaagagctcttccagacacacagtatgcact 4440
Db 4577 GCCCAACCCCTAGAAAGACCTGGCTGGCTGGAAGAGCTCTTCCAGACACCACTATGCAC 4636
QY 4441 gacaagcccaagcactcaagagaaaactaccaaaaatagcctgcagatcaacaacacagacca 4500

Db 4637 GACAACCCACGACTCACGAGAAAACATAGCAAAATAGCCTGCAGATCACAAACGAGACCCA 4696
QY 4501 gtggacacacccaacaaagctccaaagccacagtcctcaagagaagtcctcaggaaagtggaagta 4560
Db 4697 GTGGACACACCAACAAGCTCCAAAGCCACAGTCTCAAGAGAAGTCTCAGGAAGTGGACGTA 4756
QY 4561 gaagaagaattcttcgcactcagggaaacgaacaccctcagcaggcaagcca tgcacaca 4620
Db 4757 GAAGAAGAAATCTTCCCACTCAGGAAACGCAACACCATCAGCAGGCAAGCCATGCACACA 4816
QY 4621 cccaaaccagcagtaagtgtgagaaaaacatactacgcattttatgggaactccagtcgag 4680
Db 4817 CCCAAACCCAGCAGTAAGTGGTGAGAAAAACATCTAGCGCATTTATGGGAATCTCCAGTGCAG 4876
QY 4681 aaactggacctcacagagaacttaactggcagcaagacagcgtacaaaactcctaaggaa 4740
Db 4877 AAATGGAGCTGACACAGAACTTAACTGGCAGCAAGAGCGGTACAAACTTCTTAAGGAA 4936
QY 4741 aaggccagctctagaagacctggctgttttaaagagctcttccagacacagagctcac 4800
Db 4937 AAGGCCACGGCTCTAGAAGACCTGGCTGTAAAGAGCTCTTCCAGACACAGAGTCCAC 4996
QY 4801 actgaggaaatcaatgataacgataaaaaactgcocaaagtgcctgcataa tcttcacaacca 4860
Db 4997 ACTGAGGAATCAATGACTAACGATAAACTGCCAAAGTAGCCCTGCAAACTCTTCACAACCA 5056
QY 4861 gacctagacaaaaacccagaaagctccaaagcagcgtcacaagacatccctggggaagtg 4920
Db 5057 GACCTAGACAAAACCCAGCAAGCTCCAAGCGACGCTCAAGACATCCCTTGGGGAAGTG 5116
QY 4921 ggcgtgaagaagagagctcctcagttggcaagctcacacagacatcaggagagactaca 4980
Db 5117 GCGGTGAAAGAGAGCTCCTAGCAGTTGGCAAGCTCACACAGACATCAGAGAGACTACA 5176
QY 4981 cacacacacagcagcagcgaagagatgtaagagcatgaaagcatttatgagtcctca 5040
Db 5177 CACACACACACAGAGCAACAGAGATGGTAAGAGCATGAAAGCATTTATGGAGTCTCCA 5236
QY 5041 aagcagactcttagactcagcagcagcttaactggcagcaagcagcagctgagaactcct 5100
Db 5237 AAGCAGATCTTAGCTCAGCAGCAGTCTAAGTGGCAGCAAGAGCGCTGAGAACTCCT 5296
QY 5101 aagggaagctctgaagtcctgaagacctggcggcttcactcagagctcttccagacacca 5160
Db 5297 AAGGGAAGCTCTGAAGTCCCTGAAGACCTGGCGGCTTCATCGAGCTCTTCCAGACACCA 5356
QY 5161 agtcacactaaggaatcaatgactaatgaaaaactaccaaagtatcctcacagagcttca 5220
Db 5357 AGTCACACTAAGGAATCAATGAACTAATGAAAACCTTACCAAAAGTATCCTTACAGAGCTTCA 5416
QY 5221 cagccagacctagtgacacaccccaacaaagctccaagccacagcccaagagagctccag 5280
Db 5417 CAGCCAGACCTAGTGGACACCCCAACAAGCTCCAAGCCACAGCCCAAGAGAGACTCAGG 5476
QY 5281 aaagcagacactgaagaagaatttttagcatttagaatacaaacgcctacagcagcaaa 5340
Db 5477 AAAGCAGACACTGAAGAAGATTTTATAGCATTTAGGAAACAAGGCCATCAGCAGGCAAA 5536
QY 5341 gccatgcacacaccccaaacccagcagtaggtgaaggaagaagacatacaacgcgtttttggga 5400
Db 5537 GCATGTCAACACCCCAACCCAGCAGTAGTGAAGGAAGAAAGACATCAACACTGTTTTGGGA 5596
QY 5401 actccagtgcaaaaactggaccagccaggaatttaccctggcagcaatagacgctcacaa 5460
Db 5597 ACTCCAGTCAGAAACTGGACAGCCAGGAAATTTACCTGGCAGCAATAGAGCGCTACAA 5656
QY 5461 actcgtgaaggaaaagcccgagctctcagaagaactgactggcttcagagagcttttccag 5520
Db 5657 ACTCGTAAGGAAAAGGCCCAGGCTCTAGAAGAACTGACTGGCTTCAGAGAGAGCTTTCCAG 5716
QY 5521 acaacatgcactgataaacccccacagctgtagagaaactacccaaaaatactctctgcaaa 5580

Db 5717 ACACCATGCATGATAACCCACAGCTGATGAGAAAATACCAAAAAAATACTGTGCAA 5776
Qy tctcgcgaatcagaccagcgagcaccccaacaaacaaacagcaacgcccacgaagaagc 5640
Db 5777 TCTCGCAATCAGACCCAGCGACACCCCAACAACAACAAAGCAGCGCCCAAGAGAAGC 5836
Qy 5641 ctcaagaaagcagcgtatagagaagaatttttagcatctcaggaactaacccatcagca 5700
Db 5837 CTCAGAAAGCAGAGCTAGAGAGAAGAAATTTTAGCATTTACAGAAACTAACCATACGCA 5896
Qy 5701 ggcgaagccatgcacacgcctaaagcagcagtaggtgaagagaagacatacaacattt 5760
Db 5897 GGCRAAGCCATGCACACGCCCTAAAGCAGCAGTAGGTGAAGAGAAAGACATCAACACATTT 5956
Qy 5761 gtggggactccagtggagaactggacctgtcgtagaaatttacctggcagcaagagacgg 5820
Db 5957 GTGGGACTCCAGTGGAGAAACTGGACCTGCTAGGAATTTACCTGGCAGCAAGAGACGG 6016
Qy 5821 ccacaaactcctaagaaagcccaaggctctagaagatctggcttcaagagctc 5880
Db 6017 CCACAAACTCCTAAAGAAAAGGCCAAGGCTCTAGAAGATCTGGCTGGCTTCAAAAGAGCTC 6076
Qy 5881 ttccagacaccaggtccacactgaggaatcaatgaccgatgacaaaatcacagaagtatcc 5940
Db 6077 TTCAGACACCCAGGTCACTGAGGAATCAATGACCGGATGACAAATCACAGAAGTATCC 6136
Qy 5941 tgcgaatctccacaaacagaccagtcacaaaccccaacaaagctccaaagcaacgactcaag 6000
Db 6137 TGCRAATCTCCACACACAGACCCAGTCAAAACCCCAACAAGCTCCAACCAAGACTCAAG 6196
Qy 6001 atatccttggggaagttaggtgtgaagaaagaggttccctaccagctggcgaactcacacag 6060
Db 6197 ATATCTTTGGGGAAGTAGGTGTGAAGAAAGAGGTCTTACCAGTCGGCAAGCTCACACAG 6256
Qy 6061 acgtcaggaagacacacacacacacagacagagacagcagagatgaaagagcatcaaa 6120
Db 6257 ACGTCAGGGAAGACACACACACACAGAGAGAGCAGAGAGATGGAAGAGCATCAAA 6316
Qy 6121 gcgtttaggaattctgcaagcagatgctggaccacagcaaaactatggaaactggatggag 6180
Db 6317 GCGTTTAGGAATCTGCAACACAGATGCTGGACCCAGCAAACTATGGAACCTGGGATGGAG 6376
Qy 6181 aggtggccaaagacaccttaaggaagagcccaatcactagaagacctggccggcttcaaa 6240
Db 6377 AGGTGGCCAAAGAACCTTAAGGAAGAGGCCCAATCACTAGAAAGACCTGGCGGGCTTCAAA 6436
Qy 6241 gagcttccagacacacagacacactgaggaatcaacaactgatgacaaaactacaaa 6300
Db 6437 GAGCTCTTCCAGACACCAAGACACACTGAGGAATCAACAACTGATGACAAAACCTACCAAA 6496
Qy 6301 atagcctgcaaatctccaccaccagaaatcaatggacactccaaacagcagagagggcg 6360
Db 6497 ATAGCTGCAAAATCTCCACACACAGAAATCANTGGACACTTCCAACAAGCACAAGGAGCGG 6556
Qy 6361 cccaaaacacacttbgggaaagggatagtggaagagctctcagccctgaagcagctc 6420
Db 6557 CCAAAAACACCTTTGGGGAAGGGATATAGTGGAAAGAGCTCTCAGCCCTTGAAGCAGCTC 6616
Qy 6421 acacagaccacacacacagacaaagtaccagagatgaggaataaaggatcaacagtgttc 6480
Db 6617 ACACAGACCAACACACAGACAAAGTACCAGGATAGGATAAAGGCATCAACGTGTTC 6676
Qy 6481 agggaaactgcaaacagaaactggaccagcagcaagtgtlaactggtagcaagagggcag 6540
Db 6677 AGGGAACCTGCNAACAGAAACTGGACCCACGACCAAGTGTAACTGGTAGCAAGAGGCAG 6736
Qy 6541 ccaagaaactcctaaggggaaaaagcccaaccccttagaagacttggcttggcttgaagagctc 6600
Db 6737 CCAAGAACTCCTTAAGGGAAGAAAGCCCAACCCCTTAGAAGACTTTGGCTGGCTTGAAGAGCTC 6796
Qy 6601 ttccagacaccagtatgactgacaagcccacgactcacgagaaaactacaaaatagcc 6660
Db 6797 TTCCAGACACCCAGTATGCATGACAAGCCCCAGACTCACGAGAAAACCTACCAAAAATAGCC 6856

Qy 6661 tgcagatctccacaaccagagaccagtggtgtaaccccaaaatcttcaagccacagtcacaag 6720
Db 6857 TGCAGATCTCCACACACACAGACCCAGTGGGTACCCCAACAATCTTCAAGCCACAGTCCRAAG 6916
Qy 6721 agaaactcagaagaaagcagaogtagagaagaatccttagcactcaggaaaaagcaacacca 6780
Db 6917 AGAAGTCTCAGAAAGCAGACGCTAGAGGAAGAATCTTTAGCACTCAGGAAAGCAACACCA 6976
Qy 6781 tcaataggaagaactatgacacaccccaaacacagcagaggtgatgaaaaacacatgaaa 6840
Db 6977 TCAGTAGGAAGACTATGACACACCCCAACACAGCAGGAGGTGATGAGAAACACATGAAA 7036
Qy 6841 gcatttatggaaactccagtcagaaaattggacctggcaggaaaatttacctggcagcaaa 6900
Db 7037 GCATTTATGGAACTCCAGTGCAGAAATTTGGACCTGCCAGGAAATTTACCTGGCAGCAA 7096
Qy 6901 agatggccacaaactcctaaggaagaaagccaggctctagaagacctggctggcttcaaa 6960
Db 7097 AGATGGCCACAACCTCCTAAGGAAAAGGCCACAGGCTCTAGAAAGACTGGCTGGCTTCAAA 7156
Qy 6961 gagctctccagacacacagcagcactgacaaagccacagcactgatgaaaaactacaaaata 7020
Db 7157 GAGCTCTTCCAGACACACAGGCACTGACAAAGCCCACTGATGAGAAAACCTACCAAAATA 7216
Qy 7021 gctgcaaatctccacaaccagacccagtggaacacccacagcaagcacaagcaacggccc 7080
Db 7217 GCCTCAAAATCTCCACACACAGACCCAGTGGACACCCCAAGCAAGCAACAGCGCC 7276
Qy 7081 aagagaacctcagaagaaagcagacgtataggaagaatttttagcactcaggaaaaagaca 7140
Db 7277 AAGAGAAACCTCAGGAAAAGCAGACGCTAGAGGAAGAATTTTATAGCACTCAGGAAAGCAACA 7336
Qy 7141 ccacagcaggaacaaagccatggacaccccaaacacagcagtaagtatgaaaaaatc 7200
Db 7337 CCATCAGCAGGCAAAAGCCATGGACACCCCAACCAAGCAGTAAAGTGTAGAGAAAATATC 7396
Qy 7201 aacacatttggaaactccagtcagaaaactggacctgctaggaataattacctggcagc 7260
Db 7397 AACACATTTGTGGAACCTCCAGTGCAGAACTGGACCTGCTAGGAANAATTTACCTGGCAGC 7456
Qy 7261 aagagacgcacacagactccttaaggaagagctgaggctctagaggacctggtgcttc 7320
Db 7457 AAGAGACACCCACACACTCTTAAGGAAAAGGCTGAGGCTCTAGAGACCTGGTTGGCTTC 7516
Qy 7321 aaagaactctccagacacccaggtcacactaggaatcaatgactgatgacaaaatcaca 7380
Db 7517 AAAGAACTCTTCCAGACACACAGGTCACTGAGGAATCAATGACTGATGACAAAATCACA 7576
Qy 7381 gaagtatcctgttaaactccacagccagtgcatcctcaaacctcaagaagctcccaagcaa 7440
Db 7577 GAAGTATCTCTTAATCTCCACAGCCAGAGTCAATTCAAAACCTCAAGAGCTCCAAAGCAA 7636
Qy 7441 aggtcagaataccctctggtaagtgacataagaagaagccctagcagtcagcaag 7500
Db 7637 AGGCTCAAGATACCCCTTGTAAGTGGACATGAAAGAGAGAGCCCTAGCAGTACGCAAG 7696
Qy 7501 ctcacaggaactcagggagaactacgaacacacacagacgacgaagagatagtaag 7560
Db 7697 CTCACCGGACATCAGGGGAGACTACGCAACACACAGAGCCCAACGAGGATAGTAAG 7756
Qy 7561 agcataaagcgtttaagagagtcctccaaagcagatccttgagcccccagcagagtgtaact 7620
Db 7757 AGCATCAAAAGCGTTTAAAGAGTCTCCAAGAGAGATCTTGGACCCAGCAGCAAGTGAAT 7816
Qy 7621 ggtagcaggaagcagctgagaactcgttaaggaagagcccgctgctctagaagacctggtt 7680
Db 7817 GGTAGCAGGAGCAGCTGAGAACTCTGTAAGGAAAAGGCCGCTCTAGAGACCTGGTT 7876
Qy 7681 gacttcaagagctctctcagcaccaggtcacactgaagagatcaatgactattacaaa 7740
Db 7877 GACTTCAAGAGCTCTTCTCAGCACAGCTCACACTGAAGAGTCAATGACTATTGACAAA 7936

Qy	7741	aaacaaaaattccctgc	aaattccccccaccagaa	acttaacagac	cttgcacacgacaca	7800
Db	7937	AACACAAAAATTCCT	GCAATCTCCCAACCA	AACTAACAGACAC	TTGCCACGACGACA	7996
Qy	7801	aagagatgcccaagaca	cagctcccagaaaaaag	taaaagaggag	agctctcagcaggttgag	7860
Db	7997	AAGAGATGCCCAAGAC	ACCTCCCGAAGAAGT	AAAAAGAGAGCTCT	CTCAGCAATTTGAG	8056
Qy	7861	aggctcagcgaacata	caggggcaaaagcacaca	cacacaca	aaagaacacgcaagcggctgat	7920
Db	8057	AGGCTCAGCGCAACAT	CAGGGCAAGCACACAC	CACACACAAAGACAC	AGCAAGCGGTGAT	8116
Qy	7921	gagggatcaaaagtatt	tgaaagcaacgtgcaaa	agaagaacaccaa	ccccagtagaagaagaa	7980
Db	8117	GAGGCACTCAAAAT	TGTAAGAGCAACGT	GCAAGAACAAACCA	CACTAGTAAGAGGAA	8176
Qy	7981	cccagcaggaagaagg	caaagagcacctaaag	aaaaaaggccccaa	cccccttggaagaacctggcc	8040
Db	8177	CCCAGCAGGAAGGCC	CAAGACCACTTAAG	AAAAAGGCCCAAC	CCCTCGGAAGACCTGGCC	8236
Qy	8041	ggcttcacagagctct	gtgaacatacaggtcac	actcacaggaatac	actgactgctggtgcaaa	8100
Db	8237	GGCTTCACAGAGCTCT	GTGAACAATCAGGTCA	CTCAGGAACTCACT	CACTGCTGTGGCANA	8296
Qy	8101	gocactaaaaatccct	gcggaattcccccaact	agaagtgtgtagacac	caacagcaagcaca	8160
Db	8297	GCCACTAAAAATACC	TCTCGCAATCTCCCA	CTAGAACTGGTAGAC	ACCACAGCAAGCACA	8356
Qy	8161	aagaggcatctcaggaca	cgctgcaagaagggtaca	aggtaaaaaagagcct	cttcagcagctc	8220
Db	8357	AAGAGGCAATCTCAG	GACACGTGTGCAGA	AGGTACAGTAAAA	AGAAAGACGCTTCA	8416
Qy	8221	aagttcacacaaacata	caggggaaacccaggtg	cagacacaagaac	cagcaggtggaagat	8280
Db	8417	ANGTTCACAAACATCT	CAGGGGAAACCCAG	GATGCAGACAAGAA	ACCAGCAGAGGTGAAGAT	8476
Qy	8281	aaagcatcaaaagcat	tgaaggaaattgtgcaaa	acagacacccggtc	ccagcagcaagtgtta	8340
Db	8477	NAAGGCATCAAGCAT	TGAAGGAATCTGCA	AAACAGACACCG	SGCTCCAGCAGCAAGTGA	8536
Qy	8341	actggcagcaggaagac	ggccaaagagcacccag	ggaaaaagtgccca	agccatagaagaacctta	8400
Db	8537	ACTGGCAGCAGGAGC	CGGCCAAGACACCC	AGGAAAGTGCCCA	AGCCATAGAAGACCTTA	8596
Qy	8401	gctggcttcaaaagacc	cagcagcagctcacact	gaagaatacaatgact	gtatgacacaaacc	8460
Db	8597	CCTGGCTTCAAAAGAC	CCACGACGAGGTCA	CTGAAGAACTCAAT	GACTGATGACAAAAACC	8656
Qy	8461	actaaaatccctgc	aaattcatccacagaa	actagaagacacccg	caacagctcaaaagata	8520
Db	8657	ACTAAAATACCCTG	CAATCATCACAGAAC	TAGAAGACACCGCA	AAACAGCTCAAGAGATA	8716
Qy	8521	cggcccagacacgtg	cccaagaagtagaagt	gaaggaggagctgt	tagcagtttgccaag	8580
Db	8717	CGGCCACAGACACGT	CGCCCAAGAAGT	AGAAGTGAAGAG	AGGAGCTGTTAGCAGT	8776
Qy	8581	ctcacacaaacctcag	gggagaccagcacccg	acacaaagacccg	taggtaggggcaaaa	8640
Db	8777	CTCACAAAACCTCAG	GGGGAGACCGCAC	CCGACCAAGAGCC	CGGTAGGTGAGGGGCAAA	8936
Qy	8641	ggcagaaagcatttaa	gcaaacctgcnaagc	ggaaagctggacgc	agagaagttaattggc	8700
Db	8837	GGCAGAAAGCATTT	AAGCAACCTGCMA	AGCGGAACGTGAG	CGCAGAGAGATGTA	8896
Qy	8701	agcagagacagccaag	agcaactaaggaaaa	ggccccaa	cccccttggaagacctggccagc	8760
Db	8897	AGCAGGAGACAGCCA	AGAGCACCCTAAG	AAAAAGGCCCAAC	CCCTGGAAAGACCTGGCC	8956
Qy	8761	ttcaaaagctctctaa	acacacagccacact	gaggaactggc	caaaatgggtgctgctgat	8820
Db	8957	TTCAAGAGCTCTCT	CAAAACACCAAG	CCCACTAGGAA	CACTGGCAAAATGGTGT	9016
Qy	8821	agctttcaagcgtctc	aaagcaaacacacct	gcagctgga	aaaaacctcttaaaata	8880

[illegible]

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE	1 (sites)								
AUTHORS	Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G., Horrigan,S., Soppet,D.R. and Weaver,Z.								
TITLE	Cancer gene determination and therapeutic screening using signature gene sets								
JOURNAL	Patent: WO 0194629-A 7459 13-DEC-2001; Avalon Pharmaceuticals (US)								
FEATURES	Location/Qualifiers								
	1..12515								
BASE COUNT	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
ORIGIN	4166 a	3048 c	2928 g	2373 t					
	Query Match 99.5%; Score 9715; DB 6; Length 12515;								
	Best Local Similarity 100.0%; Pred. No. 0;								
	Matches 9765; Conservative 0; Mismatches 1; Indels 0; Gaps 0;								
QY	1	atgtggccacgagacgtcgttactatcaaaaggagcggggtcgacggtcccccacttt	60						
DB	197	ATGTGGCCACGACGAGCGCTGGTTACTATCAAAAGGAGCGGGTCGACGCTCCCACTTT	256						
QY	61	ccctgagcctcagcacctgctgttttgaagggtattgaaatgacatccgtatccag	120						
DB	257	CCCTCGAGCCTCAGCACCTGCTGTTTGGAAAGGGGTATTGAATGTGACATCCGTATCCAG	316						
QY	121	cttctcgttgtgcacaaacaacattgcacagttgaaatccatgagcaggaggaataatta	180						
DB	317	CTTCTCTGTGTGTCAAACAACACATTCGCAAAATTTGAAATCCATGACGAGGAGGCATATT	376						
QY	181	cataatttcagttccaaaatccacaacagtaaatgggtctgttattgatgagcctgta	240						
DB	377	CATAATTTTCAGTTCACAAATCCAAACACAGTAATAGGGTCTGTTATTGATGAGCCTGTA	436						
QY	241	cggctaaacatggagatgataaactattattgatcgttcccttcaggtatgaaaatgaa	300						
DB	437	CGGCTAAAACATGGAGATGTAATAACTATTATTGATCGTCTTCCAGGTATGAAAAATGAA	496						
QY	301	agttctcagaatgaaaggaagtaactgaatttccagaaaaatacgtgaacagagagcca	360						
DB	497	AGTCTTCAGAAATGGAAGGAAGTCAACTGAATTTCCAAGAAAAATACGTGAACAGGAGCCA	556						
QY	361	gcagctgctgtcacaagatcagcttctcttctgacctgatgagaaagctcaagattcc	420						
DB	557	GCAGCTGCTGTCTCAAGATCTAGCTTCTCTCTGACCCCTGATGAGAAAGCTCAAGATTCC	616						
QY	421	aaggcctattcaaaaatcactgaagaaaaagtttcaggaaaaatccctcagggtacatatcaag	480						
DB	617	AAGGCTATTTCAAAATCACTGAAGGAAAAGTTTCAGGAAATCCTCAGGTACATATCAAG	676						
QY	481	aatgtcaaaagaagcagtagccgcagatgactcaaaagacagtggtgtgtcagggaacaact	540						
DB	677	AATGTCAAAAGAAGACAGTACCGCAGATGACTCAAAAGACAGTGTGCTCAGGGAACAAC	736						
QY	541	aatgttcattcctcagaacatgctggacgtaattggcagaaatgcagctgattccattct	600						
DB	737	AATGTTCAATTCCTCAGAAACATGCTGGACGTAAATGGCAGAAATGCAGCTGATCCCATTTCT	796						
QY	601	ggggattttaagaataattccagcgttaaatagtagccgttatggagaataagagct	660						
DB	797	GGGGATTTTAAAGAAATTTCCAGCGTTAAATTAATGAGCCGTTATGGAGAAATTTGAAGTCT	856						
QY	661	gttcccaactacacaatgcttgcacaatagcaaaaaaatgaatccctcttggaaagctt	720						
DB	857	GTTCCCACTACACAATGCTCTTGACAATAGCAAAAAAATGAATCTCCCTTTTGAAGCTT	916						
QY	721	tatgagtcagtgaaagagagttggatgataatcacaaaaagaaaaatgtctctacagtat	780						
DB	917	TATGAGTCAGTGAAGAAGAGTTGGATGTTAAATTCACAAAAGAAAATGTCTCTACAGTAT	976						

Qy	781	tgtagaaaatctggattacaaaactgattacgcaacagagaaaaaagtgtgatggttta	840
Db	977	TGTAGAAAATCTGGATTACAAAATGATTACGCAACAGAGAGAAAGTGTCTGATGTTTA	1036
Qy	841	cagggggagacccaactgttggtctcggttaagtcaagaccaaaatctgtgtggagcgcg	900
Db	1037	CAGGGGGAGAGCCAACTGTTGGTCTCGCGTAAGTCAAGACCAGAAATCTGTGTGGAGCGC	1096
Qy	901	cacgctgtggcagagcctgttcacctgaacaagagcttgaccagaaacaaggggaaggga	960
Db	1097	CACGCTGTGGCAGAGCCCTGCTTCACCTGAACAAGAGCTTGACCAGAACAGGGGGAAGGA	1156
Qy	961	agagcgtggagctgtgttcagactcccagcaaggctgtggcgccagctttctctctat	1020
Db	1157	AGAGACGTGGAGCTGTGTTTCAGACTCCAGCAAGGCTGTGGGCGCCAGCTTTCTCTCTAT	1216
Qy	1021	gagccggttaaaaatgaagacccctgtacaataattcacagcaacaaaattctccacaaaaa	1080
Db	1217	GAGCCGGCTAAATGAAGACCCCTGTACAATATTACAGCAACAAAATTTCTCCACAAAAA	1276
Qy	1081	cataagaacaaagacctgtatactactgttagaagagaatctgtgaatctgggttaaaagt	1140
Db	1277	CATAAGACAAAGACCTGTATACTACTGTTAGTGAAGAGAAATCTGTGAATCTGGGTAAAGT	1336
Qy	1141	gaaggcttcaaggctgtgtgataaaaaactcttactcccaggaagctttcaactagaaatoga	1200
Db	1337	GAAGGCTTCAAGGCTGGTGTATAAAATCTTACTCCAGCAAGCTTTCAACTAGAAATCGA	1396
Qy	1201	acaccagctaaagttgaagatgcagctgactcgtccactaagcaagaaaatctctctcc	1260
Db	1397	ACACCAGCTAAAGTTGAAGATGCAGCTGACTTGCCACTAAGCCAGAAAATCTCTCTCC	1456
Qy	1261	aaaaccagaggaagtattctctcacagatgtggaagtctcgcctacggaactgaattcac	1320
Db	1457	AAAACCAGAGGAAGTATTCCTACAGATGTGGNAAGTTTCGCCCTACGGAACCTGAATTCAC	1516
Qy	1321	aatgagccatttttaactctgtggtcactcaagttagagaggaagatccaaaagattcc	1380
Db	1517	AATGAGCCATTTTAACTCTGTGGCTCACTCAAGTTGAGAGGAAGATCCAAAAGGATTCC	1576
Qy	1381	ctcagcaagcctggaaaaattgggcaactcagctgacagagatgtgctctgggttaactggt	1440
Db	1577	CTCAGCAAGCCTGAGAAATTTGGGCACACTACGCTGCACAGATGTCTCTGGGTACTGGT	1636
Qy	1441	cttatttcagttgatcatcaacaactttggtgattccattaaatgagagtggaggaatacct	1500
Db	1637	CTTAGTTCAGTTGATATCAACAACCTTTGGTGATTCCATTAAATGAGAGTGAGGGAATACCT	1696
Qy	1501	tgaagaagagggcgtgtgtccctttgtggcgacacctaaagacctgaactatttgatgaaaac	1560
Db	1697	TTGAAAAGAAAGCGGTGTGTCCTTTGGTGGCACCTTAAGACCTGAACTATTTGATGAAAAC	1756
Qy	1561	ttgctctctaatacgcctctcaaaaggagagagcccaacaaagaagaagctctgtgta	1620
Db	1757	TTGCTCTCTAATACGCCCTCTCAAAAGGGGAGAGGCCCAACCAAGAAAGTCTCTGGTA	1816
Qy	1621	atgcaacatccacctgtcctgaagaaaaatcatcaagggaacagcctcaaccatcaggaataa	1680
Db	1817	ATGCACACTCCACCTGTCTCTGAAGAAATCATCAGGAACAGCCCTCAACCATCAGGAAA	1876
Qy	1681	caagagtcaggttcagaaaatccatgtggaagtgaaggcacaaaagcttggttataagcct	1740
Db	1877	CAAGAGTCAGGTTTCAGAAATCCATGTGGAAGTGAAGGACAAAAGCTTGGTTATAAGCCCT	1936
Qy	1741	ccagctctagctcttaggaaaaactccagttgccagtgatcaacgcgtaggtccctgcataa	1800
Db	1937	CCAGCTCTCTAGTCTTAGGAAAACACTCCAGTTGCCAGTGATCAACGCCGTAGGTCTTCGCAA	1996
Qy	1801	acagccctctgtctccagcagcaaatctcagacagaggttccctaagagagagaggaagaa	1860
Db	1997	ACAGCCCTCTGCTTCAGCAGGCAAAATCTCAGACAGAGGTTCTTAAGAGAGGAGGACAAAGA	2056
Qy	1861	gtggcaacctgccttcaaaagagagtgctctatcagccgaagtcaacatgatattttacag	1920

Db 2057 GTGGCAACCTGCCTTCAAAAGAGAGTGTCTATCAGCGGAGTCAACATGATATTTTACAG 2116
QY 1921 atgatatgtccaaagaagaagtgtgcttcggaagcaaatctgattgtgcaaatca 1980
Db 2117 ATGATATGTTCAAAAGAGAAGTGTGCTTCGAGAGCAAACTGTATGTGTCAAAATCA 2176
QY 1981 tgggcagatgtagtaaaacttggtagcaaaacaaacaaactaaagtcatataaacatggt 2040
Db 2177 TGGGCAGATGTAGTAAACTTTGGTGTCAAAACAAACAAACTAAAGTCAATAAACATGGT 2236
QY 2041 cctcaaggtcaatgaacaaaggcaaaagaaagacctgtactctccaaagaagcgtgtgggc 2100
Db 2237 CCTCAAGGTCAATGAACAAAGGCAAAAGGCAAAAGAACCTGTACTTCCAAAGAAAGCCTGTGGGC 2296
QY 2101 gaagtccacagtcgaatttagtacaggccacgcgaacctctccttgtaccataataatagg 2160
Db 2297 GAAGTTCACAGTCAATTTAGTACAGCGCCACGCAAACTCTCCTTGTACCATATAATAGGG 2356
QY 2161 aagctcactgaaaaagtacatgtgcctgtcgcacctacagagtgtctcaacaacttc 2220
Db 2357 AAAGCTCATACTGAAAAAGTACATGTGCCTGTCTGACCCCTACAGAGTGTCAACAACCTTC 2416
QY 2221 atttccaaacaaaatggacttlaaggagaatctcttcagggaatagctgaaatgttcaag 2280
Db 2417 ATTTCCAACCCAAAATGGACTTTTAAGGAAGATCTTTTCAGGAATAGCTGAAATGTTTCAAG 2476
QY 2281 accccagtgaagagcaaccgcagtttgacaagcacatgtccacatcgctatttcaaatcca 2340
Db 2477 ACCCCAGTGAAGGAGCAACCCGCAAGTTGACAAAGCACATGTACATTCGCTATTTCAAAATTC 2536
QY 2341 gagaaattgcttgaaaaacagtttcaaggaaactgattcaggagaagaacctctgctcccc 2400
Db 2537 GAGAAATTGCTTGGAAAAACAGTTTCAAGGAACGTGATTCAGGAGAAGCAACTCTGCTCCCC 2596
QY 2401 acctcagaagtttgaggaaatgtgttcttcagtgacagaaatgcagcaaaaacagcca 2460
Db 2597 ACCTCAGAGAGTTTTCGGAGGAATGTGTCTTCAGTGCACAGAATGCAGCAAAACAGCCA 2656
QY 2461 tctgataaatgtcctgcaagccctcccttaagacgagctgatttagagaaaatggaac 2520
Db 2657 TCTGATAAATGCTCTCGAAGCCCTCCCTTAAGACGCGAGTGTATTAGAGAAATGGAAAC 2716
QY 2521 tagcaaaaaccccaggaaacctcaaaaatgactctctctggagacaaaactctcagat 2580
Db 2717 GTAGCAAAAACGCCAGGAACACCTACAAAATGACTTCTCTGGAGACAAAACCTTCAGAT 2776
QY 2581 actgagacagagccttcaaaaacagtatccactgtaaaacaggtcaggaaggtctacagag 2640
Db 2777 ACTGAGACAGAGCCTTCAAAAACAGTATCCACTGTAAACAGGTCAGGAAGGTCTACAGAG 2836
QY 2641 ttcaggaatatcacagaagctactctgtggaagtgaagtgaaagaacaaatcacagaaatt 2700
Db 2837 TTCAGGAATATACAGAAGCTACCTGTGGAAAGTTAAGAGTGAAGAAACAATACAGAAATT 2896
QY 2701 gttgagtcactctaaaaagaggtcagagggcaacactactacaaacaaaggagagaagga 2760
Db 2897 GTTGTGTGATCCTTAAAAAGAGGTCAAGAGGCAAGGCAACACTACTACAACAAAGGAGAGAAGGA 2956
QY 2761 gagatgaaggaatagaagaccttttgacatatataaggaataattgaaattaaagaa 2820
Db 2957 GAGATGAAGGAATAAGAAAGACCTTTTGTGACATATTAAGGAATAATTTGAATTTAAAGAA 3016
QY 2821 aacga tgaagaatgaagcaatgaagagatcaagaaccttggggcagagaaatgtgcacca 2880
Db 3017 AACGATGAAGAAGATGAACCAATGAAGAGATCAAGAACTTGGGGCAGAGAAATGTGCACCA 3076
QY 2881 atgtctgacctgacagacctcaagagcttgcctgtatcacagaactcatgaagacagcca 2940
Db 3077 ATGCTGTGACCTGACAGACCTCAAGAGCTTGGCCTGTATACAGAACTCATGAAAGACACGGCA 3136
QY 2941 cgtggccagaatctcctccaaaccgaagatcatgtccaaggccaccaagagtgaagaagc 3000

Db 3137 CGTGGCCAGAAATCTCTCCAAACCCCAAGATCATGTCCAAAGCACCAAGAGTGTGAAAGGC 3196
QY 3001 aaaaactcactaaaatgccttgcagtcattacaaccagaacaaataaaacaccccaacacac 3060
Db 3197 AAATCAGTCAAAATGCCCTGCCAGTCATTACAACCCAGAACCAATAAACACCCCAACACAC 3256
QY 3061 aaaaaaacacagltgaagagcatcccttgggaaaagttaggtgtgaaaagaagagctccttagca 3120
Db 3257 AAAAAACAACAGTTGAAGGCATCCCTGGGAAAGTAGGTGTGTAAGAAAGAGCTCCTTAGCA 3316
QY 3121 gtccggaagtccaacagcagctcaggggagacacacgacacacgacagagagcagcagga 3180
Db 3317 GTCGGCAAGTTTACACGGAGCTCAGGGGAGACACACGACACAGAGAGCCAGAGGA 3376
QY 3181 gatggcaagagcatcagaaccttaaggaggtctccaaagcagatccttgagaccagcagcc 3240
Db 3377 GATGGCAAGAGCATCAGAACGTTTAAGGAGTCTCCAAAGCAGATCTCTGGACCCAGAGCC 3436
QY 3241 cgtgtaacttggaaatgaagaagtyggccaagaaccttaagaaagagggcccgagtcactagaa 3300
Db 3437 CGTGTAACTGGAATGAAGAAGTGGCCAAGACGCCTAAGGAAGAGCGCCAGTCACTAGAA 3496
QY 3301 gacctgtgtggcttcaagagctcttccagacacacaggtccctctgagggaatcaatgact 3360
Db 3497 GACCTGGCTGGCTTCAAGAGAGCTTTTCCAGACACACAGGTCCTCTGAGGAATCAATGACT 3556
QY 3361 gatgagaaactaccaaaatagcctgcgaatctccaccaccagaaatcagtggaactcca 3420
Db 3557 GATGAGAAAACCTACCAAAATAGCCTGCAAAATCTCCACCACCAGAAATCAGTGGACACTCCA 3616
QY 3421 acaagcacaagaacaaatggccttaagagaagtcttcaggaaaagcagatgttagagggaagattc 3480
Db 3617 ACAAGCACAAAGCAATGGCTTAAGAGAAGTCTCAGGAAAGCAGATGTAGAGGAAGAATTC 3676
QY 3481 tttagcactcaggaanaactaacaccatcagcagggaaagccatgctttagccccaacacagca 3540
Db 3677 TTAGCACTCAGGAAACTAACACCATCAGCAGGGAAGCCATGCTTACGCCCAAAACAGCA 3736
QY 3541 ggagggtgagagaagacatataagcatatttgggaactccagtcagaaaacttggacctg 3600
Db 3737 GGAGGTGATGAGAAAGACATTTAAAGCATTTATGGGAATCTCAGTGCAGAAACTGGACCTG 3796
QY 3601 gaggaaacttacctggcgcaaaagacaggtctacagactccttaaggaaaaagggcccggt 3660
Db 3797 GCAGGAACCTTACCTGGCAGCAAAAGACAGCTACAGACTCTTAAGGAAAGGGCCAGGCT 3856
QY 3661 ctagaagacctgctgtgcttttaaagagctcttccagactccttggtcacaccaggaattta 3720
Db 3857 CTAGAAGAGCCTGGCTGGCTTTTAAAGAGCTCTTCCAGACTCCTTGGTCACACCGGGAATTA 3916
QY 3721 gtggctgtgtgtaaaaccactaaaa taccctgcgactctccacagtcagacccagtgagac 3780
Db 3917 GTGGCTGTGTGTTAAACCACTTAAATACCTTGCAGCTCTCCACAGTCCAGACCCAGTGGAC 3976
QY 3781 acccaacaagaacaaagcaacgacccagagaagtatacaggaagaagcagatgtagaggga 3840
Db 3977 ACCCCAACAAGCACAAAGCAACGACCCCAAGAGAAGTATCAGGAAAGCAGATGTGTAGAGGA 4036
QY 3841 gaaccttagcgtgcaggaatctaattgccatcagcagggcaaaagccatgcacagccctaaa 3900
Db 4037 GAACCTTTAGCTGTGAGGAATCTAATGCCATCAGCAGGCAAAAGCCATGCACAGCCCTAAA 4096
QY 3901 ccactagtaggtgaagagaaaagacatcatatttggggaactccagtcagagaactg 3960
Db 4097 CCATCAGTAGGTGAAGAGAAAGACATCATATTTTGTGGGAACCTCAGTGCAGAAACTG 4156
QY 3961 gacctgacagagaacttaaccggcagcagagagcgccacaactccttaagaagaaggcc 4020
Db 4157 GACCTGACAGAGAACTTAACCGCGCAGCAAGACGCGCCACAAACTCCTTAAGGAAGAGGCC 4216
QY 4021 caggtcttgaagacctgactgtctttaagagctcttccagaccctcttcatactgaa 4080
Db 4217 CAGGCTCTGGAAGACCTGACTGGCTTTTAAAGAGCTCTTCCAGACCCCTGTGTATCATGAA 4276

Qy 4081 gaagcagtggtgctggcaaaactactaaaaatgccctgcggaattcttccaccagaatca 4140
Dy 4277 GAAGCAGTGGCTGCTGGCAAAACTACTAAATATGCCCTCGGAATCTTCTCCACCAGAAATCA 4336
Qy 4141 gcagacaccccaacaagacacaaagcagcagcccaagacacatttggagaaaaaggaacta 4200
Dy 4337 GCAGACACCCCAACAAGCACAAAGACAGCCCAAGACACCTTTGGAGAAAAGGACGTA 4396
Qy 4201 cagaagagctctcagccctgaagaactcacagacatcaggggaaaccacacacaca 4260
Dy 4397 CAGAAGGAGCTCTCAGCCCTGAAGAAGCTCACAGACATCAGGGGAAACCACACACACA 4456
Qy 4261 gataaagtaccagaggtgaggaataaagcatcaacgcgtttaaggaaactggcaaaacag 4320
Dy 4457 GATAAAGTACCAGGAGGTGAGGATAAAGCATCAACGGCTTTAGGGAAACTGCAAAACAG 4516
Qy 4321 aaactggaccagcagcaagtgttaactggttagcaagaggaccccaaaaactaaagaaaag 4380
Dy 4517 AAATGGACCCAGCAGCAAGTGTAACTGGTATGCAAGAGGCCACCCAAAACCTAAAGGAAAAG 4576
Qy 4381 gcccaaccccttagaagacctggtggtggaagagctcttccagacaccagtgact 4440
Dy 4577 GCCCAACCCCTTAGAAGACCTGGCTGGCTGGAAAGAGCTCTTCCAGACACCAAGTATGCACT 4636
Qy 4441 gacaagccacgactcacgagaataactaccaaaatagcctgcagatcacaaaccagaccca 4500
Dy 4637 GACAAGCCACGACTCACGAGAAAACCTACCAAAATAGCCTGCAGATCACAAACGACCCA 4696
Qy 4501 gtgacacacacaagctccaagccacagctccaagagaagctcctcaggaagtggaacta 4560
Dy 4697 GTGACACACCAACAAGCTCCAAGCCACAGTCCAAGAGAAAGTCTCAGGAAAAGTGACGTA 4756
Qy 4561 gaagaagaattctcgactcagaaacgaacacacatcagcagcaagccatcacaca 4620
Dy 4757 GAAGAAGAATTCTTCGCACTCAGGAACGAACACCATCAGCAGGCAAGGCCATGCACACA 4816
Qy 4621 cccaaaccagcagtaagtgggtgagaaaaacatctacgcatttatgggaactccagtgag 4680
Dy 4817 CCCAAACCAGCAGTAAGTGGTGAGAAAAACATCTAGCATTATATGGGAACCTCCAGTGCA 4876
Qy 4681 aaactggacctgcagagaacttaactggcagcaggaagcagcgtctacaaaactcctaagaa 4740
Dy 4877 AAATGGACCTGCAGAGAACTTAACCTGGCAGCAAGAGACGGCTACAAAACCTCTAAGGAA 4936
Qy 4741 aagccacagctctagaagacctggctggttttaaagagctcttccagacacaggtcac 4800
Dy 4937 AAGCCCAAGGCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACACAGAGGTCA 4996
Qy 4801 actgaggaatcaatgactaacgataaaaactgccaaagttagcctgcaaatcttcaaccca 4860
Dy 4997 ACTGAGGAATCAATGACTAACGATATAAACTGCCAAAGTAGCCTGCAAAATCTTCACAACA 5056
Qy 4861 gacctagacaaaaccccaagcctccaagcagcagcgtccaagacatcctctgggaaagtgc 4920
Dy 5057 GACCTAGACAAAACCCAGCAAGCTCCAAGCGACGGCTCAAGACATCCCTGGGAAAAGTG 5116
Qy 4921 ggcgtgaaagagagctcttagcagttggcaagctccacagacatcagggagagactaca 4980
Dy 5117 GCGGTGAAAGAGAGCTCTTAGCAGTTGGCAAGCTCACACAGACATCAGGAGAGACTACA 5176
Qy 4981 cacacacacagagcccaacaggagatggttaagatgaaagcatttatggagtctcca 5040
Dy 5177 CACACACACAGAGCCCAACAGGAGATGTTAAGAGCATGAAGCATTTATGAGTCTCCA 5236
Qy 5041 aagcagacttagactcagcgaagcttaactggcagcaggaagcagctgagaactcct 5100
Dy 5237 AAGCAGATCTTAGACTCAGCAGCAAGTCTAACTGGCAGCAGAGGCGAGCTGAGAACTCCT 5296
Qy 5101 aaggaaagctcgaagctccatgaagacctggcggcttcatcgagctcttcagacacaca 5160
Dy 5297 AAGGAAAAGTCTGAAGTCCCTGAAGACCTGGCCGGCTTCATCGAGCTCTTCAGACACA 5356

Qy 5161 agtcacataagggaatcaatgactaatgaaaaaactacaaaagtatcctctacagagcttca 5220
Dy 5357 AGTCACACTAAGGAATCAATGACTAATGAAAAAACTACCAAAGTATCTCTACAGAGCTTCA 5416
Qy 5221 cagccagaccttagtgacacaccccaacagctccaagccaagcccaagaagaagctcaggg 5280
Dy 5417 CAGCCAGACCTTAGTGGACACCCCAACAGCTCCAAGCCACAGCCCAAGAGAAGTCTCAGG 5476
Qy 5281 aaagcagacactgaagaagaatttttagcatttaggaacaaacgcctacagcaggaaca 5340
Dy 5477 AAAGCAGACACTGAAGAAGAAATTTTAGCATTTAGGAACAAACGCCATCAGCAGCCAAA 5536
Qy 5341 gccatgcacacaccccaaccagcagtagtggaagagaagacatacaacgctttttggga 5400
Dy 5537 GCCATGCACACACCCCAACCCAGCAGTAGGTGAAGAGAAGAAACATCAACACGTTTTTGGGA 5596
Qy 5401 actccagtcgagaaactggaccagcaggaatttacctggcagcaaatagacggctacaa 5460
Dy 5597 ACTCCAGTGCAGAAACTGGACCAGCCAGGAAATTTACCTGGCAGCAATAGACGGCTACAA 5656
Qy 5461 actcgtaaaggaaaaagcccagcgtctagaagaactgactggcttcagagagcttttccag 5520
Dy 5657 ACTCTTAGGAAAAGGCCAGGCTCTAGAAAGAACTGACTGGCTTCAGAGAGCTTTTCCAG 5716
Qy 5521 acaccatgactgataaccccaacagctgatgagaaaactaccaaaaaatactctgcaaa 5580
Dy 5717 ACACCATGACTGATAACCCACACAGCTGATGAAAAACTACCAAAAAATATCTCTCAAA 5776
Qy 5581 tctccgcaatcagacccagcggagaccccaacaaacacacaaagcggcccaagagaagc 5640
Dy 5777 TCTCCGCAATCAGACCCAGCGGACCCCAACAAACACAAAGCAACGGCCCAAGAGAAGC 5836
Qy 5641 ctcaagaagcagcagctagaggaagaatttttagcatcagagaactaacaacatcagca 5700
Dy 5837 CTCAGAAAGCAGAGCTAGAGGAAGAAATTTTAGCATTTCCAGAAACTAACACCATCAGCA 5896
Qy 5701 ggcaagccatgcacacgcctaaagcagcagtaggtgaagagaagacatacaacacattt 5760
Dy 5897 GSCAAAGCCATGCACACGCCTAAAGCAGCAGTAGGTGAAGAGAAAAGACATCAACACATT 5956
Qy 5761 gtggggaactcagtgagaaaaactggacctgctaggaaaatttacctggcagcagagaagcgg 5820
Dy 5957 GTGGGAGCTCCAGTGGAGAAACTGGACCTGCTAGAAAATTTTACCTGGCAGCAAGAGACGG 6016
Qy 5821 ccacaaactcctaaagaaaaagccagcgtctagaagatctggtggtctcaagagctc 5880
Dy 6017 CCACAAACTCTTAAGAAAAAGCCAGGCTCTAGAAAGATCTGGCTGGCTTCAAGAGACTC 6076
Qy 5881 ttccagacacccaggtcactgaggaatcaatgacgatgacaaaatcacagaagtatcc 5940
Dy 6077 TTCCAGACACCCAGGTCACACTGAGGAATCAATGACCGATGACAAAATCACAGAAGTATCC 6136
Qy 5941 tgcataatctcaacacccagacccagtcataaaccccaaaagctccaagcaacgactcaag 6000
Dy 6137 TGCAAAATCTCCACACACCAGACCCAGCTCAAAAACCCCAACAAAGTCCAAGCAACGACTCAAG 6196
Qy 6001 atatccttggggaagagtggtgtaagaagaggttctctaccagtcggcgaagctcacacag 6060
Dy 6197 ATATCTTGGGGAAGTAGTGTGAAAAGAAAGAGGTCTCTACCAGTCGGCAAGCTCACACAG 6256
Qy 6061 acgtcagggaagaccacacagacacagagagcagcaggagatggaagagcatcaaa 6120
Dy 6257 AGCTCAGGGAAGCACACACAGACACAGAGACAGCAGGAGATGGAAGAGCAGCATCAAA 6316
Qy 6121 gcgtttaagggaatctgcataagcagatgctggaccacgaacaaataggaaactgggagtgag 6180
Dy 6317 GCGTTTAAAGGAATCTGCAAAAGCAGATGCTGGACCCCAAGAACTATGGAATGGGATGGAG 6376
Qy 6181 aggtggcaagaacacctaaagaagagcccaatcactagaagacctgcccggcttcaaa 6240
Dy 6377 AGGTGGCCNAGAACACCTTAAGNAGAGGCCCAATCACTAGAGACCTGCGCCGGCTTCAAA 6436
Qy 6241 gagctcttcagacacaccagaccacactgagggaatcaaacactgtgacaaaaactacaaa 6300

D	b	6437	 GAGCTCTTCCAGACACACAGACACACTGAGGAATCAACAAGTGTATGACAAAACCTACCGAAA	6496
Q	y	6301	atagcctgcgaattctccacacagaatacaatggacactccaaacagcacaagagagcg	6360
D	b	6497	ATAGCTTGCAAAATCTCCACCACAGANTCAATGGACATCTCAACAAAGCAGGAGGCGG	6556
Q	y	6361	cccaaaacacctttggggaaaaggdataagtggaagagctctcagccctgaagcagctc	6420
D	b	6557	CCCAAAACACCTTTGGGAAAAGGGATATAGTGAAGAGCTCTCAGGCCCTGAACGACGCTC	6616
Q	y	6421	acacagaccacacacagacaaagatccaggagatgaggataaaggatcaacagtgttc	6480
D	b	6617	ACACAGACCACACACACAGACAAAGTACCAGGAGATGAGGATAAAGGCATCAACGTGTC	6676
Q	y	6481	agggaaactgcaaacacagaaactgaccaccagcaagtgttaactggtgacaaagagcag	6540
D	b	6677	AGGAAACTGCAAAACAGAAACTGGACCCAGCAGCAAGTGAACCTGATGTCAGCAAGGCAG	6736
Q	y	6541	ccaagaactcctaaggggaaaagcccaacccctagaagacttggctggcttgaagaagctc	6600
D	b	6737	CCRAGAACTCCTTAAGGGAAAAGCCCAACCCCTAGAAAGACTTGGCTGGCTTGAAGAGCTC	6796
Q	y	6601	tccagacaccagtaatgactgacaagcccagactcacgagaaaactaccaaatagcc	6660
D	b	6797	TTCCAGACACCACTATGCACTGCAAGGCCAGCCAGCTACGAGAAAACTACCAAAATAGCC	6856
Q	y	6661	tgagatctccacaacacagaccagctgggtaccaccaaatcttcaagcaacagccaag	6720
D	b	6857	TGCAAGATCTCCACACACAGACCCAGCTGGGTACCCCAACAATCTTCAAGCCACAGTCCAAAG	6916
Q	y	6721	agaagtctcaggaaaagcagctagaggaagaatccttagcactcaggaaacagcaacca	6780
D	b	6917	AGAAGTCTCAGGAAAACGACGCTAGAGGAAGAATCCTTAGCACTCAGGAACGAAACCA	6976
Q	y	6781	tcagtgaggaaaagctatggacacaccccaacacagcaggaggtgatgagaagacatgaaa	6840
D	b	6977	TCAGTAGGAAAAGCTATGGACACACCCAAACAGCAGCAGGCTGATGAGAAAGACATGAAA	7036
Q	y	6841	gcattatggaaactccagtgcagaaattggactgcagaaaatttacctggcagcaaa	6900
D	b	7037	GCATTTATGGGAACTCCAGTGCAGAAATTTGGACCTGCCAGGAAATTTACCTGGCAGGCAAA	7096
Q	y	6901	agatggccacaactcctaaggaagggcccaggctctagaagacctggctggctcaaa	6960
D	b	7097	AGATGGCCACAAATCCTTAAGGAAAAGGCCAGGCTCTAGAAGACCTGGCTGGCTTCAAA	7156
Q	y	6961	gagctcttcagacaccaggcactgacaagcccacgcgactgatgaaaaactaccaaaata	7020
D	b	7157	GAGCTCTTCCAGACACACAGGCATGCAAGGCCACGACTGATGAGAAAACTACCAAAATA	7216
Q	y	7021	gcctgcgaatctccacaacacagaccagctggacccccacagcaagcacaagcaagccc	7080
D	b	7217	GCCTGCAAAATCTCCACACACACCCAGTGGACACCCAGCAAGCAACAGCAACGGCCC	7276
Q	y	7081	aagagaaacctcaggaaaagcagctagaggaagaatttttagcactcaggaaacgaaca	7140
D	b	7277	AAGAGAACTCAGGAAAACGACGCTAGAGGGAAGAATTTTATAGCACTCAGGAACGAAACA	7336
Q	y	7141	ccatcagcaggcaagccatgtgacaccccccaaacacagcagtaagtgtgagaaaaatc	7200
D	b	7337	CCATCAGCAGGCAAAAGCCATGGACACCCCAACCAACGACAGTAAAGTGTAGAAAAATATC	7396
Q	y	7201	aacacattgtggaactccagtgcagaaactggaacctcttaggaaatttacctggcagc	7260
D	b	7397	AACACATTTGTGGAAACTCCAGTGCAGAAACTTGGACCTTCTAGGAAATTTACCTGGCAGC	7456
Q	y	7261	aagagacagccacagactcctaaggaaaagcctagggtctctagaggacctgggtggcttc	7320
D	b	7457	AAGAGACGCCACAGACTCCTTAAGGAAAAGGCTGAGGCTCTTAGAGGACCTGGTTGGCTTC	7516
Q	y	7321	aagaactcttcagacacacaggtcacactgaggaatcaatgtactgatgacaaaaatcaca	7380

Db	7517	AAAGAACTCTTCCAGACACACAGGTCCACACTGAGGAATCAATGACTGTATGACAAAATCACA	7517
Qy	7381	gaagtatccctgttaattctcacagccagagtcattcaaaaacctcaagaagctccaagcaa	7440
Db	7577	GAAGTATCCTGTAAATCTCCACAGCCAGAGTCAATTCAAAACCTCAAGAAGCTCCACAGCA	7636
Qy	7441	aggtctcaagataccctcgttgaaagtggagatgaaagaagagcccttagcagtcagcaag	7500
Db	7637	AGGCTCAAGATACCCCTGTGTAAAGTGGACATGAAGAAGAGGCCCTTAGCAGTCAGCAAG	7696
Qy	7501	ctcacacggacatacagggagactacgcaaaacacacacagagcccaacaggagagatagtaag	7560
Db	7697	CTCACAGCGACATACGGGGAGACTACGCAAAACACACACAGAGGCCAACAGGAGATAGTAAG	7756
Qy	7561	agcatcaaaagcgtttaagagtcctcaaagcagatcctcggaccagagcaagtgtaact	7620
Db	7757	AGCATCAAAAGCGTTTAAAGGAGTCTCCAAAGCAGATCTCTGGACCCAGCAGCAAGTGTAACT	7816
Qy	7621	ggtagcagggagcagctggaactcgtaaaggaagggcccgctgctctagaagacactggtt	7680
Db	7817	GGTAGCAGGAGGCAGCTGAAACTCTGTAGGAAAGGCCCGTCTCTAGAAACACTGGTT	7876
Qy	7681	gacttcaaaagagctcttctcagcaacaggtctacactgaaggtcaatgactattgacaaa	7740
Db	7877	GACTTCAAGAGGTCTTCTCAGCACCAAGTCTACACTGAAGAGTCAATGACTATTTGACAAA	7936
Qy	7741	aacacaaaattcccigcaaatctcccaccagaaactaaacagacactgcacacagacaca	7800
Db	7937	ANCAAAAATTCCTGTCAANTCTCCCCACCAAGAACTAACAGACACTGCCACGAGCACA	7996
Qy	7801	aagagatgccccaagacacgctcccaggaagaagttaaagaaggagctctctcagcagttgag	7860
Db	7997	AAGAGATGCCCAAGACACGTCCACGGAAGAAGTAAAGAAGGAGCTCTCAGCAGTTGAG	8056
Qy	7861	aggtctacgcaaacatacagggcaagacacacacacacacaaagaacccagcgaagcggtag	7920
Db	8057	AGGCTCACGCAAAACATCAGGGCAAGACACACACACACAAAGAACCACAGCAAGCGGTGAT	8116
Qy	7921	gagggcatcaaaagtattgaagcaacgtcgaagagaagaacccaaacccagtagaagagaaa	7980
Db	8117	GAGGGCATCAAGTATTGAAACNACGTGCAAGAGAAGAAACCAACCCAGTTAGAAAGAGAA	8176
Qy	7981	cccagcaggaagaaggccaagagcacctaaggaaaaaggcccaacccctggaagacactggcc	8040
Db	8177	CCCAGCAGGAGAAGGCCACAGCACCTAAGGAAAAGGCCCAACCCCTGGAAGACCTGGCC	8236
Qy	8041	ggcttcacagagctctctgaacaatacaggtcacactcagggaaatactagctcgtcggcaaa	8100
Db	8237	GGCTTTCACAGAGTCTCTGAAACATCAGGTCTACACTCAGGAATCACTGACTGTCTGGCAAA	8296
Qy	8101	gcacataaataccctgcgaattctcccaccactagaagtggtagacaccacacagcagacaca	8160
Db	8297	GCCACTAAAATACCTTGGCAANTCTCCCCACCTAGAAAGTGGTAGACACCACAGCAAGCACA	8356
Qy	8161	aagaggcatctcaggacacgctgcagaaggTacaagTaaaagaagagccttctcagcagctc	8220
Db	8357	AAGAGGCATCTCAGGACACGTGTGCAGAGAGTGTCAAGTAAAGAAGAGAGCCTTTCAGCAGTC	8416
Qy	8221	aagttcacacaaacatacaggggaaccacacggaTgcagacaaagaacccagcaggTgaagat	8280
Db	8417	AAGTTTACACAAAACATCAGGGGAACACCCGGATTCGACACAAGAANAACAGCAGGTGAAGAT	8476
Qy	8281	aaaggcatcaaacattgaaggaattctgcaaacagacacccggtccagcagcgaagtga	8340
Db	8477	AAAGGCATCAAAAGCATTTGAAGGAANTCTGCAANAACAGACACCCGGCTTCCAGCAGCAAGTGT	8536
Qy	8341	actggcagcaggagacggccaagagcacccagggaagTgcccgaagccaTagaagaccta	8400
Db	8537	ACTGSCAGCAGGAGAGCGCCAGAGCACCCAGGAAAGTGGCCCAAGCCATAGAAGACCTA	8596
Qy	8401	gctggcttcaaaagaccacagcaggtcacactgaagaaatcnaatgactgtaTgacaaacc	8460
Db	8597	GCTGGCTTCAAAAGCCAGCAGCAGGTACACTGAAGAACTCAATGACTGTATGACAAAAC	8656

Qy 8461 actaaataccctgcaaatcatcaccagaactagaaagacacaccgcaacagctccaagaga 8520
|||||
Db 8657 ACTAAATACCTGCAAAATCATCCAGAACTAGAAAGACACCCGCAACAGCTCAAGAGA 8716
Qy 8521 cggcccaagacagctgcccagaaatagaaagtgaaggaggagctgttagcagctggcaag 8580
|||||
Db 8717 CGGCCAGGACACGTCGCCAGAAATAGAAAGTGAAGGAGGAGCTGTTACGAGTTGGCAAG 8776
Qy 8581 ctcacacaaacctaggggagaccacgcacaccgacaaagagccgttaggtgagggcaaa 8640
|||||
Db 8777 CTCACACAAACCTCAGGGAGNCCACGCACACCCGACAAAGAGCCGTTAGGTGAGGCGCAA 8836
Qy 8641 ggcacgaagcatttaagcaacctgcaagcggaacgtggacgcgaagaagatgtaattggc 8700
|||||
Db 8837 GGCACGAAAGCATTTAAGCAACCTCAAAGCGGAACGTTGGACGCAGAAATGTAATTGCG 8896
Qy 8701 agcaggagacagcaagagcacctaaaggagaaagcccaacccctggaagacctggccagc 8760
Db 8897 AGCAGGAGACGCCAAGAGCACCTTAAGGAAAGGCCCAACCCCTGGAAGACCTGGCCACG 8956
Qy 8761 ttccaagagctctccaacacacagccacactgaggaactggcgaatggctgctgctaat 8820
|||||
Db 8957 TTCCAAGAGCTCTCTCAACACACGACGCCACACTGAGGAACTGGCAAAATGGTCTGCTGAT 9016
Qy 8821 agctttacaagcgtctccaaagcaaacacctgacagtggaaacctctaaaaatatccaga 8880
|||||
Db 9017 AGCTTTACAAGCGCTCCAAGCAAAACACACTGACACTGTGGAACACCTTAAAAATATCCAGA 9076
Qy 8881 agagttcttgggcccttaagtagaaccctggagacgtggtgaagcaccagagacct 8940
Db 9077 AGAGTTCTTCGGGGCCCTTAAGTAGAACCCGCTGGGAGACGTTGTAAGCACCCAGACCCCT 9136
Qy 8941 gtaaatcacaaagcaaaagcaaaccttccctgccccacactgccttcaaaagggaggt 9000
Db 9137 GTAAATCACAAGCAAAAGCAACACTTCCCTGCCCCCACCCTGCTTCAAGAGGGAGGT 9196
Qy 9001 ggcaagatgaagcgtcacgggaaccagagagcgtgctgcatgccagcaccagagaa 9060
Db 9197 GSCAAAGATGGAAGCGTCACGGGAACCAAGAGCGTGCCTGTCATGCCAGCACAGAGAA 9256
Qy 9061 attgaggaggagctccagcagcaagaagcagagaggtgctccagggcagagagcaaa 9120
Db 9257 ATTGTGGAGAGCTGCCAGCAGCAAGAGCAGAGAGGGTGTCTCCAGGCGCAAGAGCAAA 9316
Qy 9121 tcattccgaaccctggtcatcatgaagaaagtttgaggactctgcgcaaaaagaattgaa 9180
Db 9317 TCATCCGAACCCGTGCTCATCATGAAGAGAACTTTGAGGACTTCTGCAGAAAAGAAATTGAA 9376
Qy 9181 cctgcggaagagctgaacagcaacgacatgaaacccaacaaagagagaaacacaaattacaa 9240
Db 9377 CCTGCGGAAGAGCTGAACAGCAACGACATGAAACCAACAAAGAGAGAACACAAATTACAA 9436
Qy 9241 gactcgttccttgaaaaaaggggaatatccctgcctccagcgcagacgccaagataagactgag 9300
Db 9437 GACTCGGTCCTGAAAATTAAGGGAATATCCCTGCGCTCCAGACGCCAAGATAAAGACTGAG 9496
Qy 9301 gcagacagcaaaataactgaggtctttgtattagcagaaaagaatagaaataaacagaaat 9360
Db 9497 GCAGAACAGCAAAATTAACCTGAGGTCTTTGTTATTAGCAGAAAAGAAATAGAAATAAACAGAAAT 9556
Qy 9361 gaaaaagagcccatgaagacacctccccagagatggacattcagaatccagatgatggagcc 9420
Db 9557 GAAAAAGAGCCCATGAAGACCTCCCCAGAGATGGACATTCAGAATCCAGATCATGGAGCC 9616
Qy 9421 cggaaacccatacttagagacaaagtctactagaaacaaagtgcttaggtctgcttaga 9480
Db 9617 CGGAAACCCATACCTTAGAGACAAAGCTCACTGAGAACAAAAGGTGCTTGAGGTCTGCTAGA 9676
Qy 9481 cagaatgagagctccagcctaaagtggcagagagagcgagggcagaaagtagtcgag 9540
Db 9677 CAGAAATGAGACTCCAGCCCTAAGGTGGCAGAGGAGCGGAGGCGCAGAAAGTCCGGAAG 9736

Qy 9541 gttctcatgcaaatcagaagggaaagagagcaggagaaattcacagctccatgtgcctg 9600
Db 9737 GTTCTCATCGAATCAGAAAGGGAAGAGAGAGCAAGAAATTCAGACTCCATGTGCCTG 9796
Qy 9601 agatcaagaaagacaaaaagccagcctgcagcaagcactttggagagcaaaatctgtgcag 9660
Db 9797 AGATCAAGAAAGACAAAAAGCCAGCCTGCAGCAAGCACATTGGAGAGCAAAATCTGTGCAG 9856
Qy 9661 agagtcaacgaggtgtcaagaggtgtgcagaaaaatccaaaaagagctgaggaacaatgtg 9720
Db 9857 AGAGTAACGCGAGGTGTCAAGAGGTGTGCAGAAAAATCCAAAGAAAGCTGAGGACAATGTG 9916
Qy 9721 tgttcaagaaataacaaccagaaagtcatagggacagtgaaagata 9766
Db 9917 TGTGTCAAGAAATAACAACCAAGACTCATAGGACAGTGAAGATA 9962

RESULT 3
HSMK167
LOCUS
DEFINITION H.sapiens mki67a mRNA (long type) for antigen of monoclonal antibody Ki-67.
ACCESSION X65550
VERSION X65550.1 GI:415818
KEYWORDS antigen; monoclonal antibody.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 12515)
AUTHORS Gerdes Fors, J.
TITLE Direct Submission
JOURNAL Submitted (11-APR-1992) J. Gerdes, Forschungsinstitut Borstel, Div. Molecular Immunology, Parkallee 22, 2061 Borstel, FRG
REMARK sequence revised by author 13-JUL-93 and 08-OCT-93
AUTHORS 2 (bases 1 to 12515)
Schluter, C., Duchrow, M., Wohlenberg, C., Becker, M.H., Key, G., Flad, H.D. and Gerdes, J.
TITLE The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins

J. Cell Biol. 123 (3), 513-522 (1993).

54043435 Location/Qualifiers
1. .12515
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10q25-qter"
/cell_line="IM9"
/clone_lib="lambda gt11"
1. .107
/number=1
108. .288
/number=2
197. .9967
/gene="mki67"
197. .9967
/gene="mki67"
/codon_start=1
/product="antigen of the monoclonal antibody Ki-67"
/protein_id="CAA46519.1"
/db_xref="GI:415819"
/db_xref="SWISS-PROT:P46013"
/translation="MWPTRLVTIKRSGVDGPHFPLSLSTCLFGRGIECDRIQLPVPV SKQCKIEIHEQAEILHNFSSNTPTQVNGSVIDEIVRLKHGDIVIILDRSFYENESL QNGKSTEFPRKIREQPARVRSRSFSDPDEKAQDSKAYKITEGVKSGPVQVHIK NYKSTADDSKDSVAQQTNTVHSHAGRNADPTISGDFKEISSVKLYRYGEL KSVPTQCIDNSKNESFPKLYESVKELDVKSNENLOYCRKSGLTDTATSKES ADLQGETOLLVSRKSRPKSGSGHVAEPASPEQELDONKKGDRDVSQVTPSKAVG ASFPLYPEAKMKTPOVSOQONSPOKHKNKDYLTTCRESVNLKSGCFKAGDKTLTP RKLSTRNTPAKVEDAADSATKPENLSSKTRGSIPTDVEVLPTETEHNEPPLTLWT QVERIKQDSUSKLTGTTAGMCSGLPGLSSVDINNFGDSINSEGLPKRRRVSF GGHLPFLDENLPPNTPILKRGEAPT KRKSLVMHTPPVLKRIKIQPOPSGQESGSE

IHYEVKAQSLVISPAPSPRKTPVASDQRRRCKTAPASSSKSQTEVPRKGGERVATC
LQKRVLSISRQHDILQMICSKRRSGASEANLIVAKSWADVVKLGAKQOTQTQVTKHGFPQ
RSMKQRPRATPKPVGVEVHSQSTFGHANSPCTIIIGKAHTEKGVHVPARPYRVLNNF
ISNQKMDGFLSGIAEMFKTPVKEOPLTSCCHIAISNSNLLGKFOGFTSGEERPL
LPTSESFGNVFPESAONAAKOPDKCSASPLLRQCI RENGNAVKTPTNTYKWTSLT
KTSDETEPSKTVSTVNSRSGSTEFNIOKLVPESSEETIEIVECILKROKATLL
QORREGEKIEIPFTYKENTELKENDKMKRNRGWQKCAPMSDLTDLUKSLPD
TELMKOTARQNLQLODQHAQAPKSEKGIKMPQCOSLOPEINTPTHTKQOLKASLG
KVGKEELAVLAKFTTSGTTHTRPAGDGKSI RTFKESPQILDPAAVPTGMKXW
PRTPKEEAQSLEDLAGFKELFQTPGSESMDEKTKIACKSPPESDVDTPTSTKOW
PKRSLRKADVEEFALRLKLTSAQKAMLT PKPAGGDEKDIKAFMGTPVQKLDLAGTL
PGSKROLOTPPKBAQALDLAGFKELFQTPGHTEELYAAGTKTKIPCDSPQSPDPVOTP
TSTKQPKRSIKRKADVEGBELLACNLMPAGKAMHTPKPSVGEEDKIIIFVGTVPVOKL
DUTENUTGSKRRPQTPKEAQALEDLTGFKELFQTPGHTEEAAVAGTKTKMPCSESP
ESADTPTSTRROPKTLEKRDVQKELSKLKTOTSGTTHTRDKVPGDGKDSINAFRE
TAKQILDPAASVTSKRHPKTKKEAQLLEDLAGWKELFQTPDKPTTHKTTK IAC
RSQDPDPTSSKPSKSLRKVDVEEFFALRKRTPSAGKAMHTPKPAVSEKNIY
AFMGCTPVOKILDLENLTGSKRLQTPKEKAQALDLAGFKELFQTPGRHTEESWINDKT
AKVACKSPDDDKNPASSKRLKTSKGVGVKEELLAVGKLTQTSGETTHHTTEPTG
DKSMKAFESPQILDSAASTGSKRQLRTPKGSSEVPEDLAGTSELFTQTPSHTKES
MYNEKTKVYSRASQPDLDVDTPTSSKQPKRSLRKADTEEEFLAFRKQTPSAGKAMHT
PKPAGVEEDKINTFLGTPVQKLDQGNLPGSNRRLOTRKEKAQALBELGFRLELFTPT
CDTNPADTKSTTKKILCKSPQSDPADPTNTKORPKRSLKADVEEFALFKLTPSA
GRAMHTPKAAGVEEKDINTFVGTVPVEKLDLGNLPGSKRRPQTPKEKAQALDLAGFK
ELFQTPGHTEESWINDKTITEVCKSPQDPVKTPTSSKQRLKISLGVGVKPEVLPVG
KLUTQSGTKTQTHRETAGDGKSIKAPKESAKOMLDPANTYGTGMERWPTPKBEAOSLE
DIAGFKELFQTPDHTDEESTDDTKTKIACKSPPPESMDTPTSTRRQTPPLGRDIVE
ELSALKQLTQTHTRDKVPGDGDKGINVRETAKQLDPAASVTGSKRQRPTRPKRAQP
LEDLAGLKFQTPVCTDKPTTHKTTK IACRSQPDVPTGTP IPKQSKRSLRKADV
EESLALRKRTPSVKAMDTPKPAGGDEKMAFGTVPVQKLDLGNLPGSKRRWQTP
KEKAQALDLAGFKELFQTPGDKPTTDEKTKIACKSPQDPVDTPTASTKQRPKNL
RKADVEEELALRKRTPSAGKAMDTPKPAVSDKKNINTFETPVQKLDLGNLPGSKR
QTPKEKAQALDLGVGFKELFQTPGHTEESMTDDKITEVCKSPQSPESFQTSRSKQ
RLKIPKLVDMKEEPLAVSKLTQTSGETTHTRTEPTGDSKSIKAPKESKQILDPAA
VTGSRQLTRKEKAQALDLVDEKELFSAPGHTEESMTIDKNTKIPCKSPPELTD
ATSTRKCPTRKVEKVELSAVERLTQTSQSTHTHKEPAGSDGEGIKVLOKAKKKP
NPVEEPPSRRRPRAPKEKAQPLEDLAGTSELSETSGHTOESLUGAKATKIPCESPLE
VDTTASTRHLRTRVQVQVKEEPSAVKFTQTSGETTHTRTEPTGAGEDGKIKALKESA
KQTPAPAAVSUTGSRRRPRAPRAQAIEDLAGFKOPAAGHTEESTDDKTKTI PKCKSS
PELEDTATSSKRRPTRAQKVEVKEELLAVGKLTQTSGETTHTRDKPEVGEKGTAFK
OPAKRVNDAEDVIGSRROPRAPKEKAQPLEDLASFOELSOTPGHTEELANGAADSFTS
APKQTPDSCKPLKISPRVLRAKVPVGVSTROPVK SOSKNTSLPLPKRGGK
DGSVGTKRLRCHMPAPEEIVBELPAKKORVAPRAKGSSEPVVIMKRSLSLPLPKRGGK
PAELNLSMKNKEHKLODSVPENKGLSLRSRODKTEASQOITEVFLARIEIN
RNEKPKMTSPMDIQNDPDGARKPIPRDKVTENRCLRSARQNESSQPKVAEESGQ
KSAKVLQMKGKGGEAGNSDMLRSRKTQSOPAASTLESKSVQRVTRSVKRCANPK
KAEDNVCVKKITTRSHRSEDI"
289. .367
/gene="mk167"
/number=3
368. .483
/gene="mk167"
/number=4
484. .550
/gene="mk167"
/number=5
561. .596
/gene="mk167"
/number=6
597. .1676
/gene="mk167"
/note="partially excluded by splicing"
/number=7
1677. .1852
/gene="mk167"
/number=8
1853. .2165
/gene="mk167"
/number=9
2166. .2284
/gene="mk167"
/number=10
2285. .2456

/gene="mk167"
/number=11
2457. .2612
/gene="mk167"
/number=12
2613. .9457
/gene="mk167"
/number=13
9458. .9901
/gene="mk167"
/number=14
9902. .12493
/number=15
polyA_signal
9994. .9999
polyA_signal
10529. .10534
polyA_signal
11237. .11242
polyA_signal
12468. .12473
polyA_site
12494
BASE COUNT 4166 a 3048 c 2928 g 2373 t
ORIGIN
Query Match 99.5%; Score 9715; DB 9; Length 12515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9765; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 atgtggccccagagacgcctgtgtactatcaaaaggagcggggtcgagcggTccccacttt 60
|||||
Db 197 ATGTGGCCACGAGAGCGCCTGGTTACTATCAAAAGGAGCGGGTTCGACGGTCCCCACTTT 256
QY 61 cccctgagcctcagcacctcgttggtaagggttatgtgaatgtgacatccgtattccag 120
|||||
Db 257 CCCTGAGCCTCAGCACCTGTTGTTGGAAGGGGTATTGAATGTGACATCCGTATCCAG 316
QY 121 ctctcgtgtgtcgaacaaacaattgcacaaagtgcacatccatcagcaggaggaata 180
|||||
Db 317 CTTCCTGTTGTGCAAAACAACATTGCAAAATTCGAAATTCATGACGAGGAGCAATATTA 376
QY 181 caLaatttcagttccacaacatccacaacagtaaaatgggtctgttattgatgagccgtga 240
|||||
Db 377 CATAATTTCAAGTTCCACAAATCCACACAAAGTAAATGGGTCTGTTATTGATGAGCCTGTA 436
QY 241 cggctaaaaacatggagatgtataactatttatgtctccttcagggtatgaaaaTgaa 300
|||||
Db 437 CGGCTAAAACATGGAGATGTAAATACTATTATTGATCGTTCCTTCAGGTATGAAAATGAA 496
QY 301 agtcttcaaaatggaaggagtcacactgaatttcccaagaaaaatacgtgaacaggagcca 360
|||||
Db 497 AGTCTTCAGAAATGGAAGGAAGTCAACTGAAATTTCCAAAGAAAAATACGTGAACAGGAGCCA 556
QY 361 gcacgtcgtgtctcaagatctagcttctcttgaccctgatgagaaagctcaagattcc 420
|||||
Db 557 GCAGCTCGTGTCTCAAGATCTAGCTTCTCTTCGACCTCATGAGAAAGCTCAAGATTCC 616
QY 421 aagcctattcaaaaatcactgaagaaaaagtttcaggaaaaatcctcagggtacatatcaag 480
|||||
Db 617 AAGGCTATTTCAAAAATCACTGAAGGAAAAAGTTTCAGGNAATCCTCAGGTACATATCAAG 676
QY 481 aatgtcaaaagaacagttaccgcagatgactcaaaagacagttgttctcagggaacaact 540
|||||
Db 677 AATGTCAAAGAAACAGTACCGCAGATGACTCAAAAGACAGTGTGCTCAGGGAACAACT 736
QY 541 aatgttcattctccagaaacatgctggcagcgtaaatggcagaaatcagctgatccattct 600
|||||
Db 737 AATGTTTCATTCCTCAGAACATGCTGGAGCTTAATGGCAGAAATGCAGCTGATCCATTTCT 796
QY 601 ggggaatttaaaagaaatttccagcgtttaaattagtgagccgttatggagaaattgaagtct 660
|||||
Db 797 GGGGATTTTAAAGAAATTTCCAGCGTTAAATTTAGTGAGCGGTTATGAGAGAAATTTGAGTCT 856
QY 661 gttccactacacaatgtcttgacaatagcaaaaaataaataatccctctttgagagctt 720
|||||
Db 857 GTTCCCACTACACAATGTCTTTGACAAATAGCAAAAAAATGAATCTCCCTTTTUGAAGCTT 916

QY 721 tatgagtcagtgaagaagagttggatgtataaatcacaaaaagaaaaatgtctcactagtat 780
|||||
Db 917 TATGAGTCAGTGAAGAAAGAGTTGGATGTAAATACAAAAAGAAATGCTCTCAGTAT 976
QY 781 tgaagaaatctggattacaactatcagcaacagagagaagaagtcgtatgtgttta 840
Db 977 TGTAGAAAATCTGGATTTACAACTGATTACGCCAACAGAGAAAGAGTCTGATGGTTTA 1036
QY 841 cagggggagacccaactgttggtctcgcgttaagtcaagaccacaaaatctggtgagcggc 900
Db 1037 CAGGGGAGACCCAACCTGTTGGTCTCGCGTAGTCAAGACCAAAATCTGGTGGGAGCGC 1096
QY 901 cagcgtctggaagagccctgtctcaactgaacaagagcttgaccagaaacaaggggagggga 960
Db 1097 CACGCTGTGGCAGAGCCCTGCTTCACTGAACAAGAGCTTGACCAGAACAAAGGGGAAGGGA 1156
QY 961 agagacgtggagctgtgttcagactcccaagcaaggctgtgggcgcagactttcctctctat 1020
Db 1157 AGAGACGTGGAGTCTGTTTCAGACTCCACAGAGGCTGTGGGCGCCAGCTTTCCTCTCTAT 1216
QY 1021 gagccgcgtaaaaagaaacccctgtacaataattcacagcaacaaaattctccacaaaaa 1080
Db 1217 GAGCCGGCTAAAAAGAACCCCTGTACAAATTTACAGCACAAATAATCTCCACAAAAA 1276
QY 1081 cataagaaacaagacctgtatactactgtgtagaagaaatctgtgaatcttggttaaaagt 1140
Db 1277 CATAAGAAACAAAGACCTGTATACTACTGTGTAGACAGAAATCTGTCAATCTGGGTAAAGT 1336
QY 1141 gaagcctcaaggcttggtgataaaactcttactcccaaggaagcttcaactagaatacga 1200
Db 1337 GAAGCCTTCAAGGCTGGTGTATAAAACTCTTACTCCAGGAAGCTTTCAACTAGAAATCGA 1396
QY 1201 acaccagctaaagtgaagatgcagctgactctgccactaagccagaaaaatctctctcc 1260
Db 1397 ACACCAGCTAAAGTTGAGATGACAGTCACTGCGCACTAAGCCAGAAAAATCTCTCTCC 1456
QY 1261 aaaaaccaggaagattctcactagatgtgaagttctgcctacgaaactgaaattcac 1320
Db 1457 AAAACACAGAGGAAGTATTCCTACAGATGTGGAAGTCTGCGCTACGGAACCTGAAATTCAC 1516
QY 1321 aatgaagcatttttaactctgtggtcactcaagttagagagaagaatccaaaaggattcc 1380
Db 1517 AATGAGCCATTTTAACTCTGTGGCTCACTCAAGTTGAGAGGAAGATCCAAAGGATTCC 1576
QY 1381 ctcaagcagctgagaaattgggcactacagctgacagatgtgctctgggttacctggt 1440
Db 1577 CTCACAGCCCTGAGAAATGGGCACCTACAGCTGGACAGATGTGCTCTGGGTACTTGGT 1636
QY 1441 cttagttcagttgatatacaaacctttggtgattccattaatgagagttagggataacct 1500
Db 1637 CTTAGTTTCAGTTGATATCAACAACCTTTGGTGATTCATTAATGAGAGTGAGGGAATACCT 1696
QY 1501 ttgaaaaagaagcgtgtgctcttgggtgggcacctaaagacctgaactattgatgaaac 1560
Db 1697 TTGAAAAGAAGCGGTGTCTCTTTGGTGGGCACCTTAAAGCCCTGAACTATTGATGAAAC 1756
QY 1561 ttgcctctataacagcctctcaaaaggggagaaagcccaacaaagaagtcctctggtta 1620
Db 1757 TTGCTCTCTAATACGCCCTCTCAAAAGGGGAGNAGCCCAACCAAAAGAAAGTCTCTGGTA 1816
QY 1621 atgcacactccacctgtcctgaagaaaaatcatcaagggaacagcctcaaccactcaggaaaa 1680
Db 1817 ATGCACACTCCACCTGTCTCTGAAGAAATATCATCAAGGAACAGCCCTCAACCATCAGGAAAA 1876
QY 1681 caagagtcaggttcagaataatccatgtggaagtgaaggcacaaaagcttggttataagccct 1740
Db 1877 CAAGAGTCAGGTTTCAGAAATCCATGTGGAAGTGAAGGCACAAAAGCTTGGTTATAAGCCCT 1936
QY 1741 ccagctcttagtccttagaagaactccagttgcccagttgatcaacgcgttaggtccctgcaaa 1800
Db 1937 CCAGCTCTTAGTCTTAGGAAAACCTCCAGTTGCCAGTGTATCAACGCCGTAGGTCCTGCAAA 1996

QY 1801 acagccctgtctccagcagcaaatctcagacagaggttctctaagagagagagagaaga 1860
Db 1997 ACAGCCCTGCTTCCAGCAGCAAAATCTCAGACAGAGGTTCTTAAGAGAGAGGAGAGAAGA 2056
QY 1861 gggcaacctgcttcaaaaagagagtgctatatcagccgaagtcaacatgatattttacag 1920
Db 2057 GTGGCAACCTGCTTCAAAAAGAGAGTGTATCAGCCGAAGTCAACATGATATTTTACAG 2116
QY 1921 atgatatgttccaaaagaagaagtggtcttcggaagcaaaatctgattgttgcaaaatca 1980
Db 2117 ATGATATGTTCAAAAGAAGAAAGTGGTCTTCGGAAGCAAAATCTGATTGTTGCAAAATCA 2176
QY 1981 tgggcagatgttagtataaacttggtcacaaacaaacaaactaaagtcataaaacatggt 2040
Db 2177 TGGGAGATGTAGTAAAACTTGGTGCAAAACAAACAACTAAAGTCAATAAAACATGGT 2236
QY 2041 cctcaaggtcaatgaacaaaaaggcaaaagaacacctgtactcccaagaagacctgtgggc 2100
Db 2237 CCTCAAAAGGTCAATGAACAAAAGGCAAAAGAACCTGTACTCCAAGAAAGCCTGTGGGC 2296
QY 2101 gaagttcacagtcgaatttagtacagccacgcgaacctctccttgtaccataataaLaggg 2160
Db 2297 GAAGTTCACAGTCAATTTAGTACAGGCCACGCAAACTCTCCTTGTACCATATAATATAGGG 2356
QY 2161 aaagctcatactgaaaaagtacatgtgcctgtcgcacctacagagtgctcaacaacttc 2220
Db 2357 AAAGCTCATACTGAAAAAGTACATGTGCTGTGCGACCTACAGAGTGTCTCAACAACCTTC 2416
QY 2221 atttccacccaaaaatggactttaagggaagatcttttcagggaatagctgaatatgttcaag 2280
Db 2417 ATTTCCAAACCAAAAATGGACTTTAAGGAAGATCTTTTCAGGAATAGCTGAAATGTTTCAAG 2476
QY 2281 acccagtgaaagagcaaccgcagttgacaagcacatgtcacatcgctctatttcaaatca 2340
Db 2477 ACCCAGTGAAGAGCAACCCGAGTTGACAGCACATGTACATCCTCTATTTCAAATTCAC 2536
QY 2341 gagaattgtctggaaaaacagtttcaaggaaactgattcaggagagaacacctctgctcccc 2400
Db 2537 GAGAAATTTGCTTTGGAAAAACAGTTTCAAGGAACCTGATTCAGGAGAGAACTCTGCTCCCC 2596
QY 2401 acctcagagagtttggaggaatgtgttcttcagtgacagaaatgcaacaaacagcca 2460
Db 2597 ACCTCAGAGAGTTTGGAGGAATGTGTCTTCAGTGCACAGAAATGCACAAACAGCCA 2656
QY 2461 tctgataaatgctctgcaagccctcccttaagacggcagtgattagagaaaaatggaaac 2520
Db 2657 TCTGATAATGCTCTGCAAGCCCTCCCTTAAAGCGCAGTGTATTAGAGAAATGGAAC 2716
QY 2521 gtacaaaaacgcccgaggaaacacctacaaaatgacttctctggagacaaaaacttcagat 2580
Db 2717 GTAGCAAAACGCCCAGGAACACCTACAAAATGACTTCTCTGGAGACAAAACTTCAGAT 2776
QY 2581 actgagacagaccttcaaaaacagtatccactgtaaacaggttcaggaaggtctacagag 2640
Db 2777 ACTGACAGAGAGCTTCAAAAACAGTATCCACTGTAAACAGGTCAGGAAGGTCACAGAG 2836
QY 2641 ttcaggaaatatacagaagctactctgtggaagtgaagtgaaagaaacaaatcacagaaatt 2700
Db 2837 TTCAGGAATATACAGAGCTACCTCTGGGAAAGTAAAGTGAAGAAACAATAACAGAAAT 2896
QY 2701 gttgagtcactctaaaaagaggtcagaaggcaacactactacaacaaaggagagagga 2760
Db 2897 GTTGAGTGCATCTCTAAAAAGAGTCAAGAGGCAACACTACTACACAAAGGAGAGAGGA 2956
QY 2761 gagatgaaggaataagaaaagaccttttgagacataataagaaaaatattgaaattaaagaa 2820
Db 2957 GAGATGAAGGAATAGAAAAGCTTTTGTAGACATATATAGGAAATATTGAATTTAAAGAA 3016
QY 2821 aacgatgaaaaagatgaagcaatgaagagatcaagaaacttggggcagaaaatgtgcacca 2880
Db 3017 AACGATGAAAAAGATGAACCAATGAAGATCAAGAACTTGGGGCGCAAAATGTGCACCA 3076
QY 2881 atgtctgacctgacagacctcaagagcttgcctgtatcacagaactcatgaagacacggca 2940

||||| 3077 ATGTCACCTGACAGACCTCAAGAGCTTGCCGTGATACAAACTCATGAAGACACGGCA 3136
QY 2941 cgtggcagaaatctctccaaacccaagatcatgccaaggcaccaaagagtggaaaagc 3000
Db 3137 CGTGGCCAGAATCTCCTCCAAACCCAAAGATCATGTCGAAGGCACCAAGAGGTGAGAAAGC 3196
QY 3001 aaaatcctaataatgcctgcagtcattataccaggaaacaaataaaacaccccaaacac 3060
Db 3197 AAAATCACTAAAATGCCCTGCCAGTCATTACAACCCAGAACAATAAACACCCCAACACAC 3256
QY 3061 acaaaacaagttgaagcatccctggggaagtagtgtgaagaagagctccttagca 3120
Db 3257 ACAAACAACAGTTGAAGGATCCCTGGGGAAGTAGGTGTGAAGAAGAGCTCCTTAGCA 3316
QY 3121 gtcggcaagttcacacgcagtcaggggagaccacgcacacgcacagagagccagcagga 3180
Db 3317 GTCGGCAAGTTCACACGCGCTCAGGGGACACCCAGCACGACAGAGAGCCAGCAGGA 3376
QY 3181 gatggcaagagcatcagaacggttaaggagtcctccaaagcagatccttggaccacagcc 3240
Db 3377 GATGGCAAGAGCATCAGAACGTTTAAAGAGTCTCCAAAGCAGATCCTTGGACCCAGCAGCC 3436
QY 3241 cgtgaactggaatgaagaagtgcgaagaacgcctaaagaaagagagcccaagtagaa 3300
Db 3437 CGTGTAACTGGAATGAAGAAGTGGCCGAAGAACGCCTAAGGAAGAGGGCCCACTCAGTAA 3496
QY 3301 gacctggctgcttcaaaagagctctccagacaccaggtccctctgaggaatcaatgact 3360
Db 3497 GACCTGGCTGCTTCAAAGAGCTTCTCCACACACCAGGTCCCTCTGAGGAAATCAATGACT 3556
QY 3361 gatgagaaactaccaaaa tagcctgcaaatctccaccaccagaaatcagtggaactcca 3420
Db 3557 GATGAGAAAATACCAAAAATAGCTGTCMAAATCTCCACCACAGAAATCAGTGGACACTCCA 3616
QY 3421 acaagcaagaagaatggcctaagaaagtcctcaggaagcagatgtagaggagaattc 3480
Db 3617 ACAAGCAAGAAGCAATGGCTTAGAAGAGTCTCAGGAAGCAGATGTAGAGGAAGAAATC 3676
QY 3481 ttgacctcaggaaactaacaccatcagcagggaaagccatgcttacgcccaaacccagca 3540
Db 3677 TTAGCACTCAGGAACCTAACACCATCAGCAGGGAAGCCATGCTTACGCCCAACACGCA 3736
QY 3541 ggaggtagagaaagacataaagcatttatgggaactccagtgcagaaacttgacctg 3600
Db 3737 GGAGGTGATGAGAAAGACATTAAGCATTTATGGGAATCTCCAGTGCAGAACTGGACCTG 3796
QY 3601 gcaggaaactttacctggcagcaaaagacagctacagactccttaagaaaagccaggct 3660
Db 3797 GCAGGAACTTTACTGGCAGCAAAACACAGCTACAGACTCCTAAGGAAGAAGGCCAGGCT 3856
QY 3661 ctgaaagacctggctggctttaagagctcttcagactcctgtgtcacaccggaggaaat 3720
Db 3857 CTAGAAGACCTGGCTGCTTTAAAGAGCTCTTCCAGACTCCTGGTCACACCCGAGGAATTA 3916
QY 3721 gtggctgtgtaaaaccactaaaaatccctggcagctctccacagtcgaacccagtgagc 3780
Db 3917 GTGGCTGTGGTAAAAACCACTAAAATACCCCTGGCACTCTCCACAGTCAGACCCAGTGGAC 3976
QY 3781 accccaagcacaagcaaacgaccaccaagagaagtatcaggaagaagcagatgtagaggga 3840
Db 3977 ACCCCAACAGCACAAGCAACGACCCAGAGAAAGTATCAGGAAGACAGATGTAGAGGGA 4036
QY 3841 gaactcttagctgcaggaaatctaattgcccatacagcaggcaaaagccatgtcacagcctaa 3900
Db 4037 GAACTCTTAGCTGTCAGGAATCTTAATGCCATCAGCAGGCAAAAGCCATGCACAGCCCTAAA 4096
QY 3901 ccaatagtaggtgaagagaaagacatacaatatttgggaaactccagtgcagaactcg 3960
Db 4097 CCATCAGTAGGTGAAGAGAAAGACATCATATTTTGTGGAACTCCAGTGCAGAAACTG 4156
QY 3961 gacctgacagagaacttaaccgagcagagagcggcccaactccttaaggaagggcc 4020
|||||

Db 4157 GACCTGACAGAGAACTTTAACGGCAGCAAGAGAGCGGCACAAAATCTCTTAAGGAAGAGGCC 4216
QY 4021 caggctctggaagacctgactggctttaagagagctcttccagacccctggfcaactagaa 4080
Db 4217 CAGGCTCTGGAAGACCTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTCATCTGAA 4276
QY 4081 gaagcagtggtctgtggcaaaactactaaatgcccctgcgaattcttctccaccagaaatca 4140
Db 4277 GAAGCAGTGGCTGCTGGCAAAACTACTATAATGCCCTGCGAATCTTCTCCACCAGAAATCA 4336
QY 4141 gcagacacccccacaagcacaaagagcagcccaagacacactttgggagaaaaagggacgta 4200
Db 4337 GCAGACACCCCAACAAGCACAAAGAGCGCCCAAGACACTTTTGGGAGAAAAGGGAGCTA 4396
QY 4201 cagaagagagctctcagccctgaagaagctcacacagacatacaggggaaaccacacacaca 4260
Db 4397 CAGAAAGAGCTCTCAGCCCTGAAGAAGCTCACACAGACATCAGGGGAAACCCACACACACA 4456
QY 4261 gataaagtaccagaggtgaggataaaagcatcaaacgcttttaggggaaactgcacaaacag 4320
Db 4457 GATAAAGTACCAGAGGTGAGGATAAAAGCATCAACGCGTTTAGGGAAACTGCAAAAACAG 4516
QY 4321 aagctggacccacagcaagcttaactggttagcaagagcagcccaaaaaactaaggaaaaag 4380
Db 4517 AAATGGACCCAGCAGCAAGTGTAACTGGTAGCAAGAGGACCCCAAAAATTAAGGAAAAG 4576
QY 4381 gcccaacccctagaagacctggctggctggaagagctcttccagacaccagtatgcact 4440
Db 4577 GCCCAACCCCTAGAGACCTGGCTGGCTGGAAAGAGCTCTTCCAGACACCAGTATGGACT 4636
QY 4441 gacaagccacgactcacagagaaaaactaccaaataagcctgcagatcacaaacacagaccca 4500
Db 4637 GACAAGCCACGACTCAGAGAAAATATCCAAAATAGSCCTGCAGATCACAACACAGACCCA 4696
QY 4501 gtgacacacccaagaagctccaagccacagctccaagagaagctctcaggaagaagtggacgta 4560
Db 4697 GTGGACACACCAACAGCTCCAAGCCACACAGTCCAAGAGAAGTCTCTCAAGAAAGTGGACGTA 4756
QY 4561 gaagagaaattcttcgacctcaggaaagaaacaccaLcagcaggcaagcatalgcacaca 4620
Db 4757 GAAGAAGAAATCTTCGCACCTCAGSAAACGAACACCATCAGCAGGCAAGCCATGCACACA 4816
QY 4621 cccaaaccagcagtaagtggtagagaaaaacatactacgcaLttatgggaaactccagtgag 4680
Db 4817 CCCAAACCAAGCAGTAAAGTGGTAGAAAAACATCTACGCATTTATGGGAACCTCCAGTGCAG 4876
QY 4681 aaactgacctgacagagaacttaactggcagcaagagcgcgtacaaactcctaaggaa 4740
Db 4877 AAATGGACCTGACAGAGAACTTAACCTGGCAGCAAGAGAGCGGTACAAATCTCTTAAGGAA 4936
QY 4741 aagggccaggtctagaagacctggctggctttaagagagctcttccagacacagaggtcac 4800
Db 4937 AAGGCCAGGCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACACAGGCTCAC 4996
QY 4801 actgaggaatactgactaacgataaaactgccaagtgccaagtgagcctgcacaaacttcacaacca 4860
Db 4997 ACTGAGGAATCAATGACTAACGATAAAACTGCCAAAAGTAGCCTGCAAAATCTTTCACAACCA 5056
QY 4861 gacctagacaaaaaccacaagctccaagcgcgcgtcaagacatcccttggggaagtg 4920
Db 5057 GACCTAGACAAAANCCAGCAAGCTCAAGCGACGGCTCAAGACATCCCTTGGGGAAGTG 5116
QY 4921 ggcgtgaaagaagagctccttagcagttggcaagctcacacagacatacagagagactaca 4980
Db 5117 GCGGTGAAAGAAGAGGCTCCTTAGCAGTTGGCAAGCTCACACAGACATCAGGAGAGACTACA 5176
QY 4981 cacacacacagagccaacagagagatgtgaagagcatgaagacattatggaggtctcca 5040
Db 5177 CACACACACAGAGCCACAGGAGATGGTAAAGAGATGAAAGCATTTATTTGGAGTCTCCA 5236
QY 5041 aagcagatcttagactcagcagcaagcttaactggcagcaagagcagctgcgaactcct 5100
Db 5237 AAGCACATCTTAGACTCAGCAGCAAGTCTAACTGGCAGCAAGAGCGAGCTGCAAACTCCT 5296

Qy 5101 aaggaaagtctgaagtcctccctgaagaaacctggccgcttctcagctcttccagacacca 5160
Dy 5297 AAGGAAAGTCTGAAGTCCCTGAAGACCTGGCCGGCTTCATCGAGCTCTCCAGACACCA 5356
Qy 5161 agtcacataaggaataatgactaataatgaataaactacccaagatctctacagagctca 5220
Dy 5357 AGTCACACTAAGGAATCAATGACTAATGAAAAAACTACCAAGTATCTCTACAGAGCTTCA 5416
Qy 5221 cagccagacctagtgagacaccccaacagctccaagccacagcccaagagaagctcagg 5280
Dy 5417 CAGCCAGACCTTAGTGAGCACCCCAACAAGCTCCAAGCCACAGCCCAAGAGAAGTCTCAGG 5476
Qy 5281 aaagcagacactgaagaagaatttttagcattttaggaacaacacgcctcagcgacaaa 5340
Dy 5477 AAAGCAGACACTGAAGAAGATTTTAGCATTTAGGAACAAACGCCATCAGCAGCCAAA 5536
Qy 5341 gccatgcacaccccaaccccgagtagtgagaagaagacatacaacacgtttttggga 5400
Dy 5537 GCCATGCAACACACCCAAACCCAGCAGTAGTGAAGAGAAGACATCAACACCGCTTTTGGGA 5596
Qy 5401 actcagtgagaaactgagaccagcaggaataatttacctggcagaataagcagctacaa 5460
Dy 5597 ACTCAGTGCAAAACTGGACCGCCAGCGAATTTTACCTGGCAGCAATAGACGGCTACAA 5656
Qy 5461 actcgtaaaggaagggccagggctctagaagaactgactggcttcagagagcttttcag 5520
Dy 5657 ACTCGTAAGGAAGAGGCCCGCCAGCTCTAGAAGAACTGACTGGCTTCAGAGAGCTTTCCAG 5716
Qy 5521 acacatgcaatataacccacagctgtagagaaaactacaaaataactctgcgaaa 5580
Dy 5717 ACACCATGCACTGATAACCCCAACAGCTGATGAGAAAATACCAAAAAAATACTCTGCAAA 5776
Qy 5581 tctcgaactcagaccagcggaacaccccaacaaacaaagcaacgcaccaagagaagc 5640
Dy 5777 TCTCGCAANTCAGACCCAGCGGACACCCCAACAAACAAAGCAACAGCCCAAGAGAAGC 5836
Qy 5641 ctcaagaagcagactagagaagaataatttttagcattcaggaaactacacccatcagca 5700
Dy 5837 CTCAGAAAGCAGACGCTAGAGGAAGAAATTTTAGCATTCAGGAAACTAACACCATCAGCA 5896
Qy 5701 ggcaaaagccatgcacacgccttaagcagcagtagtggaaggaagacatacaacacattt 5760
Dy 5897 GGCAAAAGCCATGCACACGCCCTAAAGCAGCAGTAGGTGGAAGAGAAGACATCAACACATTT 5956
Qy 5761 gtgggactccagtggaagaactgagacctgctagagaattttacctggcagcaagaagacgg 5820
Dy 5957 GTGGGACTCCAGTGGAGAACTGGACCTGCTAGGAAATTTTACCTGGCAGCAAGAGACGG 6016
Qy 5821 ccaaaactcctaaagaaagggcctctagaagatctggctggcttcaagagctc 5880
Dy 6017 CCACAAACTCTAAAGAAAGGCCCAAGGCTCTAGAAGATCTGGCTGGCTTCAAGAGCTC 6076
Qy 5881 ttccagacacacaggttcaactgaggaaatcaatgaccgatgacaaaatcacaagaagtatcc 5940
Dy 6077 TTCCAGACACCCAGGTTCACACTGAGGAATCAATGACCGGATGACAAAATCACAGAGTATCC 6136
Qy 5941 tgcgaatctcacacacagaccagtcgaaccccaacaaagctccaagcaacgactcaag 6000
Dy 6137 TGCAANTCTCCAAACACCAAGACCCAGTCAAAACCCCAACAGCTCCAAGCAGCAGCTCAAG 6196
Qy 6001 atatacttggggaagtaggtgtgaaagaagggctctaccagctggcgaagctcacacag 6060
Dy 6197 ATATCTTTGGGAAAGTAGGTGTGAAGAAGAGGTTCTTACCAGTCGGCAGCTCACACAG 6256
Qy 6061 acgtcaggaagaccacacagacacagagagacagcagagatggaagaagcattcaaa 6120
Dy 6257 ACGTCAGGGGAACACACACAGCACACAGAGAGACAGCAGGAGATGGAAAGAGCATCAAA 6316
Qy 6121 gcgtttaaggaatctgcaaaagcagatgctgacccacgacaaactatgggaactgggatgag 6180
Dy 6317 GCGTTTAAAGGAATCTGCAAAAGCAGATGCTGGACCCAGCAAACTATGGAACTGGGATGGAG 6376

Qy 6181 aggtggccaaagacacacctaagaaaggcccccactactagaagacctggccggtcttcaaa 6240
Dy 6377 AGGTGGCCAAAGAACACCTTAAGGAAGAGGCCCAATCACTAGAAAGACTGGCCGGCTTCAAA 6436
Qy 6241 gagctcttcagacacccagaccacactgagggaatcaacaactgatgacaaaactaccaaa 6300
Dy 6437 GAGCTCTTTCCAGACACCAAGACACACCTGAGGAATCAACAACTGATGACAAAACTACCAAA 6496
Qy 6301 atagctctcaaatctccaccaccagaaatcaatggacactcccaacaagcacaagagggcg 6360
Dy 6497 ATAGCTCTCAAAATCTCCACCACCAAGTCAATGGACACTTCCAACAAAGCAAGAGGCCG 6556
Qy 6361 ccaaaaaacacctttgggaaaggatatagtggaagagctctcagccctgaagcagctc 6420
Dy 6557 CCCAAACACCTTTGGGGAAGAGGATATAGTGAAGAGCTCTCAGCCCTGAAGCAGCTC 6616
Qy 6421 acacagacacacacacagacaaaagtaccagagatgaggataaaaggatcaaacgtgttc 6480
Dy 6617 ACACAGACCCACACACACAGACAAAGTACCAGGAGATGAGGATAAAGGCATCAACGTGTT 6676
Qy 6481 agggaaactgcaaaaacagaaaactggaccagcagcaagtgttaacttggttagcagagggcg 6540
Dy 6677 AGGGAACATGCAAAACAGAAACTGGACCCAGCAGCAAGTGTAACTGGTGTAGCAAGAGCGAG 6736
Qy 6541 ccaagaactcctaaaggaaaaagcccaaccccttagaagacttggctggcttgaagagctc 6600
Dy 6737 CCAAGAACTCTTAAGGGAAGAGCCCAACCCCTAGAAGACTTGGCTGGCTTTGAAGAGCTC 6796
Qy 6601 ttccagacacacgtatgactgacaaagccacagactcacgagaaaactcccaaaaatagcc 6660
Dy 6797 TTCCAGACACCAAGTATGCACTGACAAAGCCACGACTCAGGAGAAAATACCAAAAATAGCC 6856
Qy 6661 tgcagatctccacaaccagaccagctgggtaccacaacaaatctcaagccacagctccaag 6720
Dy 6857 TGCAGATCTCCACAACACAGACCCAGTGGGTACCCCAACAATCTTCAAGCCACAGTCCAAG 6916
Qy 6721 agaactctcagaaagcagacgtagaggagaatccttagcactcaggaaaacgaacacca 6780
Dy 6917 AGAAGTCTCAGGAAAGCAGACGCTAGAGGAAGAAATCCTTAGCCTCAGGAAACGAACACCA 6976
Qy 6781 tcagtaggaaagctatggacacaccccaacccagcagaggtgtagagaaaagacatgaaa 6840
Dy 6977 TCAGTAGGGAAGAGCTATGGACACACCCCAACAGCAGGAGGTGATGAGAAAAGACATGAAA 7036
Qy 6841 gcatttatgggaactccagctcagaaattggacctgcaggaataattactctggcagcaaa 6900
Dy 7037 GCATTTATGGGAACCTCCAGTGCAGAAATTTGACCTGCCAGGAATTTACTTGGCAGCAAA 7096
Qy 6901 agatggccacaaactcctaaaggaagggcccaggctctagaagacctggctggcttcaaa 6960
Dy 7097 AGATGGCCACAAACTCCTTAAGGAAAAGGCCCAAGGCTCTAGAAAGACCTGGCTTCAAA 7156
Qy 6961 gagctcttcagacacccagggaactgacaaagcccaagactgatgagaaaaactcccaaaaata 7020
Dy 7157 GAGCTCTTTCCAGACACCAAGGCATGACAAAGCCCAAGCTGATGAGAAAACTACCAAAAATA 7216
Qy 7021 gctctcaaatctcccaaacacagaccagctggacccccagccagcaagcacaagcaagggccc 7080
Dy 7217 GCCTCAAAATCTCAACACCAACAGACCCAGTGGACACCCCAAGCAAGCAACAAAGCAAGGCC 7276
Qy 7081 aagagaacactcaggaagcagacgtagaggagaattttttagcactcaggaaaacgaaca 7140
Dy 7277 AAGAGAAACCTCAGGAAAGCAGACGCTAGAGGAAGAAATTTTAGCCTCAGGAAACGAACA 7336
Qy 7141 ccatcagaggaacaaagccaatggaccccccaaacacagcagtaagtgaatgagaaaaataatc 7200
Dy 7337 CCATCAGCAGGCAAGCCATGGACACCCCAAAACAGCAGTAAAGTGTAGAAAAATATC 7396
Qy 7201 aacacatttgtgaactccagctcagaaaactggacctcaggaataattactctggcagc 7260
Dy 7397 AACACATTTGTGGAAACTCCAGTGCAGAACTGGACCTGCTAGGAAATTTTACTTGGCAGC 7456
Qy 7261 aagagacagccacagactccttaaggaaaaaggctgaggctctagagacctgagggcttgccttc 7320

Db 7457 AAGAGACGCCACAGACTCTCTAGGAAAAGGCTGAGGCTCTAGAGGACCTGGTGGCTTC 7516
QY 7321 aaagaactcttcagagacaccaggTcaactgaggaatacaatgactgatgacaaaatcaca 7380
Db 7517 AAGAACTCTTCAGAGACACCGAGTCACACTGAGGAATCAATGACTGATGACAAAATCACA 7576
QY 7381 gaagtacctgttaattccacagcagagtgatctattcaaaacctcaagaagctccaagcaa 7440
Db 7577 GAAGTATCTCTGTAATCTCCACAGCCAGAGTCATTTAAAACTTCAAGAAGCTCCAAAGCAA 7636
QY 7441 agctcaagataccctggTgaagtggacatgaagaagaagcccttagcagtcagcaag 7500
Db 7637 AGCTCAAGATACCCCTGGTGAAGTGGACATGAAGAAGAGGCCCTTAGCAGTCAGCAAG 7696
QY 7501 ctacacgagatcatcaggggagactacgcaaacacacacagagcgaacacagagagatagtaag 7560
Db 7697 CTCACAGGACATCAGGGAGACTACGCAACACACACACAGAGCCACACAGAGATAGTAAG 7756
QY 7561 agcatcaagcgTttaaggagTctccaaagcagatcctgtgaccacagcagcaagtgttaact 7620
Db 7757 AGCATCAAGCGTTTAAGGAGTCTCCAAAGCAGATCTCTGACCCAGCAGCAAGTGAAC 7816
QY 7621 ggtagcagagcagctgaactcgttaagaaaagcccgTgctcttagaagacctgggt 7680
Db 7817 GGTAGCAGGAGGCGAGCTGGAACCTCGTAGGAAAAGGCCGCTGCTCTAGAAGACCTGGT 7876
QY 7681 gacttcaagagctctctcagcaccaggTcaactgaaagTcaatgactattgacaaa 7740
Db 7877 GACTTCAAGAGCTCTCTCAGCACCAGGTCACACTGAAGAGTCAATGACTATTGACAAA 7936
QY 7741 aacacaaaattccctgcaaatctccccaccagaactaacagacactgcacagacaca 7800
Db 7937 AACACAAAATTCCTGCAATCTCCCCACCAGAACTTAACAGACACTGCCACGAGCACA 7996
QY 7801 aacagatgcccaagacagTcccagaaaagaagTaaaagaggagctctcagcagttgag 7860
Db 7997 AAGAGATGCCCCAAGACACGCTCCAGGAAAGAGTAAAAGAGGAGCTCTCAGCAGTTGAG 8056
QY 7861 aggtctacgcaaacatacagggcaagacacacacacacaaagaaacagcaagcgtgtat 7920
Db 8057 AGGCTCAGCAACATCAGGGCAAGACACACACACACACAAAGAACACGACGCGTGAT 8116
QY 7921 gagggcatcaaatgtgaagcaactgcaagaagaataccaaacctcagtagaaggaa 7980
Db 8117 GAGGGCATCAAAATATTGAAGCAACCTGCAAGAAAGAACTAAAAGAGGAGCTCTCAGCAGTTGAG 8176
QY 7981 ccagcagagaagccaagagcacctaaagaaaagcccaacccctgaaagacctggcc 8040
Db 8177 CCCAGCAGNAGGACCAAGAGCCTTAAGGAAAAGGCCCAACCCCTGGAAGACCTGGCC 8236
QY 8041 ggcttcacagagctctctgaacatcaggtcaactcaggaatacactgactgctggcaaa 8100
Db 8237 GGCTTCACAGAGCTCTGAAACATCAGGTCACTCAGATCAGSNATCACTGACTGCTGCCAAA 8296
QY 8101 gcaactaaaataccctggaattccccactagaagtggtagacaccacagcaagacaca 8160
Db 8297 GCCACTAAAATACCCCTGCGAATCTCCCCACTAGAAAGTGTAGACACCACAGCAAGCACA 8356
QY 8161 aagagcatctcaggaacagTgtgcagaaggtacaagttaaaagaagagccttcagcagtc 8220
Db 8357 AAGAGCATCTCAGGACAGCTGTGCAGAAGGTACAAGTAAAAGAAGAGCCCTTCAGCAGT 8416
QY 8221 aagttcacacaaacatacaggggaaacccaggtgcagacaaagaacacagcaggtgaagat 8280
Db 8417 AAGTTCAACAAACATCAGGGGAAACCCAGGATGCAGACAAAGAACACGAGTGAAGAT 8476
QY 8281 aaaggcatcaagcattgaggaattctgcaaaacagacaccggctccagcagcaagtgtca 8340
Db 8477 AAGGGCATCAAGCATTTGAAGGAATCTGCAAAACACACACACCCGGCTCAGCAGCAAGTGA 8536
QY 8341 actggcagcaggaagcgcgaagagcaccagggaaagTgccaagcctaagaagacctta 8400
|||||

Db 8537 ACTGGCAGCAGGAGACGGCCCAAGAGCACCCAGGGAAAAGTGCCTCCCAAGCCATAGAAACCTA 8596
QY 8401 gctggcttcaaaagacccagcagcaggtccacactgaagaaTcaatgactgatgacaaaacc 8460
Db 8597 GCTGGCTTCAAGACCCAGCAGCAGGTTCACACTGAGAATCAATGACTGATGACAAAACC 8656
QY 8461 actaaataccctgcataatcatcaccagaactagaagacaccgcaacaagtctcaaaagaga 8520
Db 8657 ACTAAAATACCTTGCAATCATCCAGAACTAGAAAGACACCCGCAACAGCTCAAGAGA 8716
QY 8521 cgccccagagacagTctccagaaagttagaagtgaaggaggagctgttagcagttggcaag 8580
Db 8717 CGGCCAGGACACGTGCCAGAAAGTAGAAGTGAAGGAGGAGCTGTTAGCAGTTGGCAAG 8776
QY 8581 ctacacaaacctcaggggagacacgacacgcgaacaaagccggTtaggtgagggcaca 8640
Db 8777 CTCACAAACCTCAGGGGAGACCACGCACACACCAAGAGCCGCTAGGTAGGGCAAA 8836
QY 8641 ggcaagaagcatTtaagcaacctgcaagcggaacgTggacgcagaagatgttaattggc 8700
Db 8837 GGCACGAAAAGCA/TTAAGCAACCTGCAAGCGGAACGTGGACGCAGAAATGTAAATTGGC 8896
QY 8701 agcagagacagccaagcagcactaagaaaaggcccaacctggaaagacctggccagc 8760
Db 8897 AGCAGGAGACAGCCAAGACCACTTAAGAAAAGGCCCAACCCCTGGAAGACCTGGCCAGC 8956
QY 8761 ttccaaagagctctcaaacacaccagggccacactgagaaactggcaaatggtgctgtgat 8820
Db 8957 TTCCAAGAGCTCTCTCAACACCAGGCCACACTGAGGAAGTGCACAAATGCTGTGTGAT 9016
QY 8821 agctttacaagcgtccaaagcaaacacacctgacagTggaaaacctctaaaaatatccaga 8880
Db 9017 AGCTTTACAAGCGCTCCAAAGCAAAACACCTGACAGTGGAAAACCTCTTAAAAATATCCAGA 9076
QY 8881 agagtctctcgggccctcaaatagtagaacccgTgggagaagTgtagaccaccagagacct 8940
Db 9077 AGAGTTCTTTGGGCCCTTAAAGTAGAACCCGTGGGAGACGTGTGTAAACACACAGAGACCT 9136
QY 8941 gTaaaatcaaaagcaaaagcaaacctTccctgcccccaactgccttcaagaggggaggt 9000
Db 9137 GTAAATCACAAAAGCAAAAGCAACACTTCCCTGCCCCCACTGCCCCTTCAAGAGGGGAGGT 9196
QY 9001 ggcaagaTggaagcgtcaagggaaccaaagagcgtgcgtgcattgcagcaccacagaggaa 9060
Db 9197 GGCAAAAGATGGAAAGCTCACGGGAACCAAGAGCTGCGCTGCATGCGCAGCACCAGAGGAA 9256
QY 9061 attgtgagagagctccagccagcaagagaggttgcctccagggcgaagacaaa 9120
Db 9257 ATTGTGAGAGAGCTGCCAGCCAGCAAGACAGAGGTTGCTCCCAGGGCAAGAGCAAA 9316
QY 9121 tcatccgaacccgtggTcatcatgaagagaagTttgaggactctgcataaaagaaattgaa 9180
Db 9317 TCATCCGAACCCGTGTGTCATCATGAAGAGAAAGTTTCAGGACTTCTGCAAAAAGAAATTGAA 9376
QY 9181 cctgcggaagagcTgaacgcaacgcacatgaaaaaccaaaagagaggaacacaaattacaa 9240
Db 9377 CCTGCGGAAGAGTGAACAGCACACGACATGAAACCAACAAAGAGGAACACAAATTTACAA 9436
QY 9241 gactcggtccctgaaataaagggaatatccctgcgtccagagcccaagataagactgag 9300
Db 9437 GACTCGGTCCCTGAAAATTAAGGGAATATCCCTGCGCTCCAGAGCCCAAGATAGACTGAG 9496
QY 9301 gcagacagcaaatactgaggtctttgtattagcagaagaatagaaataaacagaaat 9360
Db 9497 GCAGAAACAGCAAAATACTCAGGCTCTTTGTATTAGCAGAAAGAAATACAAATAACAGAAAT 9556
QY 9361 gaaagaagcccatgaagacctccccagagatggacattcagaatccagatgtagagcc 9420
Db 9557 GAAAAGAACCCATGAAGACCTCCCCAGAGATGGACATTTCAGAATCCAGATGATGAGCC 9616
QY 9421 cggaaccccatccttagagacaaaagTcactgagacaaaaggtgcttgaggtcgtctaga 9480
Db 9617 CGGAAACCCATACCTAGACACAAAGTCACTGAGACAAAGGTCCTTGAGGTCGTCTGATGA 9676
|||||

Qy	9481	cagaatgagagctccagcctaagtggcagagagagcggaggagcagaagagtgcggaag	9540
Db	9677	CAGATGAGAGCTCCAGCGCTTAAGTGGCAGAGGAGCGGAGGCGCAGAAAGTGCGAAG	9736
Qy	9541	gttctcatgcagaatcagaagaaggaagagagagagaaattcagactccatgtgcctg	9600
Db	9737	GTTCTCATGCAGAAATCAGAAAGGGAAGAGAGAGCAATTCAGACTCCATGTGCCTG	9796
Qy	9601	agatcaagaagaacaaaagccagcctgcagcaagcacttggagagcaaaatcctgtgcag	9660
Db	9797	AGATCAAGAAACAAAAGCCAGCGCTGCAGCAACACTTTGGAGAGCAAAATCTGTGCAG	9856
Qy	9661	adagtaacgcgagtggtcaagaggtgtgcagaaaaatccaaaagcgtgagacaaatgtg	9720
Db	9857	AGAGTAACGCGAGTGTCAAGAGGTGTGCAGAAAAATCCAAAGAGCTGAGGACAATGTG	9916
Qy	9721	tgtgtcaagaataataacaacagagtgatcatagggacagtgaagata	9766
Db	9917	TGTGTCAAGAAATAACAACCAAGAGTCATAGGGACAGTGAAGATA	9962
RESULT 4			
HSWK167A			
LOCUS	HSWK167A	11435 bp	mRNA
DEFINITION	H.sapiens mki67a mRNA (short type) for antigen of monoclonal antibody Ki-67.		
ACCESSION	X65551		
VERSION	X65551.1 GI:415820		
KEYWORDS	antigen; monoclonal antibody.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 11435)		
TITLE	Gerdes Fors, J.		
JOURNAL	Direct Submission		
REMARK	Submitted (11-APR-1992) J. Gerdes, Forschungsinstitut Borstel, Div. Molecular Immunology, Parkallee 22, 2061 Borstel, FRG		
AUTHORS	sequence revised by author 14-JUL-93 and 08-OCT-93		
TITLE	2 (bases 1 to 11435)		
JOURNAL	Schluter, C., Duchrow, M., Wohlenberg, C., Becker, M.H., Key, G., Flad, H.D. and Gerdes, J.		
MEDLINE	The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins		
FEATURES	J. Cell Biol. 123 (3), 513-522 (1993)		
source	94043435		
exon	Location/Qualifiers		
exon	1..11435		
gene	/organism="Homo sapiens"		
CDS	/db_xref="taxon:9606"		
	/chromosome="10q25-qter"		
	/cell_line="IM9"		
	/clone_lib="lambda gt11"		
	1..107		
	/number=1		
	108..288		
	/number=2		
	197..8887		
	/gene="mki-67"		
	197..8887		
	/gene="mki-67"		
	/codon_start=1		
	/product="antigen of the monoclonal antibody Ki-67"		
	/protein_id="CAA46520.1"		
	/db_xref="GI:415821"		
	/db_xref="SWISS-PROT:P46013"		
	/translation="MWPTRLVITIKRSGVDGPHPLSLTCLFGRGIRCDIRIOLPVV SKOCHKTEIHQEAIIHLNFSSTNPVNGSVIDEPVLRKHGDVITIDRSFYENESL QNGKRSTFPKRREDEPARVRSKSSDDESGIPPKRRRVSGFGLRPELFDEN LPPNTPILKRGAPTKRKSILVMHTPVLKIIKEQPPSGKQESGSETHVEYKAQSLVI SPPAPSPKPTVPSDQRRRSCKTAPASSKSKSQTEVPKRGGERVATCLQKRVYSIRSQH		

DI LQIMICSKRRSGASEANLIVAKSWADVVVLGAKATOTKVIKHGQPSRMNKRORPAT
PKIPVGVHSQSTGHANSPTIIIGKATEKHVHPARVYRVNLFINQKMDKDFEVL
SGIAEMFTYVKEQPOLFTSTCHIAISNENLLGKQFOGTGDSGEPLLPSTSSFGNVE
FSAONAAKQPSDKCSASPLRQOCIRENGNAKTPRNTYKMTSLTKTSDTETESKT
VSTYNSRGSSTEFNRNIQKLPVESKSEETNTEIVECILARGQKATILLQORREGEKMEIE
RPFYETRIENIELKENDERKAMKRWGQKCAPMSDLTDULASPDTELMKDTRAGON
LLQOTDHAAPKTSKSEKITTMPQCSLOPEPINTPHHTKQOLKASLGKVGVBELLAVG
KFTRTSGTTHHREPADGKSIIRTFKESPKQILDPAARVTGMMKWPPTPKAEASLE
DLAGFKELFOTPGSESWTDEKTKIACKSPPSVSDPTSTKWPKRSUKRAADVEE
EFLARKLITPSAGKAMLPKAGGDEKDIKAFMGTPVOKLIDLAGTLPGSKRLOLPKE
KAQLEDLAGFKELFOTPGHTEELVAAGKTTKIPCDSPQSDPVDPTSTKQPKRSIR
KADVEGELLACNLMPKSAGKAMHTPKPSVGEKDIIFVFTPVQKLDENLTGSRRR
POTPKEEAALDLMTGFKELFOTPGHTEEAAGKTTKMPCESSPEADPTSTSRQ
PKTLEKRDVQKELSALKKLTQSTGTHHTDKVGGEDKSIINAFRETAKQKLDPAASV
TKSKRHPTKDEKAQLEDLAGFKELFOTPVCTDKPTHEKTKIACRSPQDPTDPTS
SKPOSKRSIRKVDVEEFALRKTPPSAGKAMHTPKPAVSGEKNLPAFMGTVPQKLDL
TENLTGSKRRLOTPEKAQALEDLAGFKELFOTGRGTHTEESMTNDKATAVACKSQPDL
DKNPASRRRLKTSILGKVGVEELIAGKLTQSTGTHHTHTTEPTGDKSKAFMESP
KOILDSASLTGSKRRLOTPKGSVEPDLAGFELFOTPSHTKESMTNPKVTYSYR
ASQDLVDPTSSKQPKRSRKADTEEFALFRKQTPSAGKAMHTPKPAVGEEDKIN
TFLGTPVOKLDQGNLPGSNRRLOTREKAQALELTGFRLEFOTPCDNDPTADEKTT
KKILCKSPQSDPADPTNTKORPKBSLKADVEEFLAFRLITPSAGKAMHTPKAAVG
BEKDINTFVGTPEKLDLGLNLPCKRRPOTPEKAKALEDLAGFKELFOTPGHTEES
MTDDKITEVCKSPQSDPVKTPSTSSKRLKISLGKVGKVEEVLPGKLTQTSKGTQTT
HRETAGDGKSKAFESAKQMLDPANYGTGMERWPTPKEEAQSLDLAGFKELFOTP
HTDEESTTDDKTTKACKSPPPESMDTPTSTRRPKTPGLKRDIVIELSALKQLTQT
HDKVPGEDKGINVRETAKQKLPAAASVTGSKRPRTPKGAQPLEDLAKGLKELFO
TPVCTDKPTHEKTKIACRSPQDPTPTIFKPOSKRSIRKADVEESLALRRKTP
SVGRAMDTPKAGGDEKMAFMGTVPVOKLDLPGNLPGSKRWPOTPEKAQLEDLAG
FKELFOTPGTDKPTTDEKTKIACKSPQDPTDPTASTQKRNLRKADVEEELAL
RKRTPSAGKAMDTPKPAVSGDEKNTFETVPQKLDLGLNLPCKRRLOTPPEKAEAL
EDLVGFKELFOTPGHTEESMTDDKITEVCKSPQSPESFKTSRSSKQRLKIPLVKDMK
EPLALEDLVDFKELFSAPGHTESMTIDKNTKIPCKSPPELTATSTKRCPTRP
EKARALEDLVDKELFSAPGHTESMTIDKNTKIPCKSPPELTATSTKRCPTRP
RKEYEELSALVRLTQTSQSTHKEPASGDEGKIVLKQRAKKPNPVEEPSRRRP
RAPKEAQLDLELSELTSGHTQESLTAGKATKIPCESPPVLEVVDTTASTRRHL
RTRVOKVQVKEEPSAVKFTQTSGETTADKEDPAGDKGKIKALKESAKQTPAPAAVGT
SRPRPAPRESAQAIEDLAGFKDPAAGHTESMTDDKTKIPCKSPELTATSTSKR
RPRTAQKVEYKEELLAYGKLTQTSGETTHTDKEPVGEGKTKAFKQPAKRNVDADV
IGSRQPRAPKEAQLDLELSELTSGHTQESLTAGKATKIPCESPPVLEVVDTTASTRRHL
KISRVLRAKVPVGVDRDPVKSKSNSTSLPPLPFKRGGKSDSVGTCTKRLRC
MPAPEEIVLELPASKQIRVAPRARGKSEPVVIMKRSURTSAKRIEPAEELNSDMKT
NKEEHLQDSVPENKQISLRSSRDQTEAQOITVEVLAERIEINRNEKKPMKTSPE
MDIQNDPQARKPIPRDKVTENKRLSARQNESSQPKVAEESGQKSAKVLQMQKG
KGEAGNSDSMCLRSRKTKSQPAASTLESKSVQVTRSVKRCANPKKAEDNVCKKIT
TSHRSEDII"
289..367
/gene="mki-67"
/number=3
368..483
/gene="mki-67"
/number=4
484..550
/gene="mki-67"
/number=5
561..596
/gene="mki-67"
/number=6
597..772
/gene="mki-67"
/note="exon 7 is partially excluded by splicing"
/number=8
773..1085
/gene="mki-67"
/number=9
1086..1204
/gene="mki-67"
/number=10
1205..1385
/gene="mki-67"
/number=11
1386..1532
/gene="mki-67"

```

exon      /number=12
1533..8377
/genes="mki-67"
exon      /number=13
8378..8821
/genes="mki-67"
exon      /number=14
8822..11413
/number=15
8914..8919
polyA_signal
9449..9454
polyA_signal
10157..10162
polyA_signal
11388..11393
polyA_signal
11414
BASE COUNT 3802 a 2828 c 2684 g 2121 t
ORIGIN

Query Match      84.8%; Score 8286; DB 9; Length 11435;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 8286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1481 atgagagtgggaatacctttgaaagaaggcgtgtgtctcttggcgccactaagac 1540
Db 597 ATGAGAGTGAAGGAATACCTTTGAAAAGAGGCGTGTGTCTTTGGTGGGCACCTAAGAC 656

QY 1541 ctgaactattgatgaatacttgctcctctaatacgccctctcaaaaggggagaagcccaa 1600
Db 657 CTGAACATATTTGATGAATACTTGCTCTCTAATACGCCCTCTCAAAAGGGGAGAGCCCCAA 716

QY 1601 ccaaaagaagttctgtgtaatcacactccactgtctctgaagaaatcatcaaggaac 1660
Db 717 CCAAAAAGAAAGTCTCTGTTAATCACACTCCACTCTCTCTGTAAGAAATCATCAAGGAAC 776

QY 1661 agctcaaccatcagaaacaaagagtcagggttcagaaatccatgtggaagtgaagcac 1720
Db 777 AGCCTCAACCCTCAGGAAAACAAGATCAGGTTTCAGAAATCCATGTGGAAGTGAAGGCAC 836

QY 1721 aaagcttggttaaaagccctccagctctctagtctcctaggaatacctcagttgccagtgtc 1780
Db 837 AAAGCTTGGTTATAAGCCCTCCAGCTCCTTAGTCTAGGAAATCCAGTTGCCAGTGTATC 896

QY 1781 aagcgcgtagttcctgcataaaagccctcgtctccagcagcaaatctcagacagaggttc 1840
Db 897 AAGCCCGTAGGTCTGCAAAACAGCCCTGCTTCCAGCAGCAAAATCTCAGACAGAGGTTTC 956

QY 1841 ctaagagagagagaaagagtggaacctgccttcaaaagagagtgtctatcagccgaa 1900
Db 957 CTAAGAGAGAGAGAAAGAGTGGCAACCTGCTCTTCAAAAGAGAGTGTCTATCAGCGGAA 1016

QY 1901 gtcaacatgatatattacagatgatattgttccaaaagaagaagtgtgtctcggaagcaa 1960
Db 1017 GTCAACATGATATTTTACAGATGATATGTTCAAAAGAAAGAAAGTGTGTCTTCGGAAGCAA 1076

QY 1961 atctgattgtgcataatcatggcgagatgtatgataaaacttggtgcataaaacacacaaa 2020
Db 1077 ATCTGATTTGTCAAAATCATGGCGAGATGTAGTAAAACTTGGTGTCAAAAACAACACAAA 1136

QY 2021 ctaagtcataaacaatggctcctcaaggtcgaatgaacaaaaggcaagaagacgtgtcta 2080
Db 1137 CTAAGTCATAAATCATGGTCTCTCAAGGTCAATGAACAAAGCAAAAGAACCTTGCTTA 1196

QY 2081 ctccaaagaagcgtgtggcgaaagttcacagttcaacttttagtacaggccacgcgaactctc 2140
Db 1197 CTCAAAAGAGCCGTGTGGCGAAGTTTACAGTCAATTTAGTACAGGCCACGCCAAACTCTC 1256

QY 2141 ctgtaccataataatagggaagactcaactactgaaaaagtacaatgtgcctgtctcgacct 2200
Db 1257 CTTGTACCAATAATTAATAGGGAAGAGCTCATACTGAAAAAGTACATGTGCCTGTCTCGACCC 1316

QY 2201 acgagtgctccaacaacttatttccaaacaaataatgactttaagggaagatctttcag 2260
Db 1317 ACAGAGTGCTCAACACTTCATTCCACCACCAAAAATGGACTTTAAGGAAGATCTTTTCAG 1376
```

```

QY 2261 gaatagctgaatgtttcaagacccccagtgaaagagcaaccgcagttgacaaagcacatgtc 2320
Db 1377 GAATAGCTGAAATGTTTCAAGACCCCACTAGTGAAGGAGCAACCCGCACTTCAACAGCACATGTC 1436

QY 2321 acatcgctatttcaaatccagagaatttgcttggaataacagtttcaagggaactgatttcag 2380
Db 1437 ACATCGCTATTTTCAAAATTCAGAGAAATTTGCTTGGAAAACAGTTTCAAGGAACATGATTCAG 1496

QY 2381 gagaagaacctctgctcccccacactcagagagttttggaggaaatgtgtcttccagtgac 2440
Db 1497 GAGAAGAACCTCTGTCTCCCACTTCAGAGAGTTTGGAGGAAATGTGTTCTTTCAGTGCAC 1556

QY 2441 agaatgcagcaaaacagccatctgtataataatgctctctgcaagccctcccttaagacgagc 2500
Db 1557 AGAATGCAGCAAAACAGCCATCTGATAAATGCTCTCTGCAAGCCCTCCCTTAAAGACGGCAGT 1616

QY 2501 gattagagaaaaatggaacgtagcaaaaacgcccagggaacactactacaaaatgactctc 2560
Db 1617 GTATTAGAGAAAATGGAAACGTAGCAAAAACGCCAGGAACACCTTACAAAATGACTTCTC 1676

QY 2561 tggagacaaaaacttcagatactagacagagccttcaaaaacagtatccacitgtaaac 2620
Db 1677 TGGAGACAAAACCTTCAGATACTTGAGACAGAGCCTTCAAAAACAGTATCCCACTGTAAACA 1736

QY 2621 ggtcaggaaggtctacagagtttcaggaatatatacagaaagctacctgttggaagttaagagt 2680
Db 1737 GGTcAGGAAGGTCTACAGAGTTTcAGGAATATACAGAAAGCTACCTGTGTGGAAGTAAAGAGT 1796

QY 2681 aagaacaaatcacagaaattgttgatgcatctctaaaaaagggttcagaaaggcaacactac 2740
Db 1797 AAGAAAACAATACAGAAATTTGTTGAGTGCATCTCTAAAAGAGGTCAGAAAGCAACACTAC 1856

QY 2741 tacaacaaaggagagagagagatgaagaaataaagaaaccttttgagacataaagg 2800
Db 1857 TACACAAAGGAGAGAGGAGAGATGAAGGAAATAGAAAGACCTTTTTCAGACATATAAGG 1916

QY 2801 aaaaattgaattaaagaagaaacgatgaaagaatgaagcaatgaagagatcagaacact 2860
Db 1917 AAAATATTGAATTTAAAGAAAACGATGAAAGATGAAAGCAATCAAGAGATCAAGAACTT 1976

QY 2861 gggggcagaaaaatgtgcacaaatgtctgacctgacagacactcaagagcttgcctgatacag 2920
Db 1977 GGGGGCAGAAATGTGACCAATGTCTGACCTGACAGACCTTCAAGAGCTTGCCTTGATACAG 2036

QY 2921 aactcatgaagacacgcgcagtcgagcaaatctctccaaacccaagatcatgccaagg 2980
Db 2037 AACTCATGAAGACACGGCACGTGGCCAGAAATCTCTCCAAACCCCAAGATCATGCCCAGG 2096

QY 2981 caccaaagagtgaagaagcgaatactactaaaaatgcctgtccagtcattacaaccagaac 3040
Db 2097 CACCAAAAGAGTGAGAAAGGCAAAATCACTAAAATGCCCTGCCAGTCATTTACAACCAAGAAC 2156

QY 3041 caataaaaccccccaacacacacacacacacacacacacacacacacacacacacacacac 3100
Db 2157 CAATAAACACCCCAACACACACACACACACACACACACACACACACACACACACACACACAC 2216

QY 3101 tgaagaagaagctccctagcagtcgcgaagtctcacgcgcagtcgagggagacacacacaca 3160
Db 2217 TGAACAAAGAGTCTCTTAGCAGTCGGCAAGTTTACACGGAGCTCAGGGGAGACCCACGACA 2276

QY 3161 cgcacagagagccagcagagagatgagcaagagatcagaacgttttaaggaggtctccaaagc 3220
Db 2277 CGCACAGAGAGCCAGCAGGAGATGGCAAGAGCATCAGAAGCTTTAAAGAGTCTTCCAAAGC 2336

QY 3221 agatctcgaccacagcagcccggtgtaactggaatgaagaagtggccaagaacgccttaagg 3280
Db 2337 AGATCTTGGACCCAGCAGCCCGTGTAACTTGAATGAAGAAAGTGGCCAAAGAACGCCCTAAGG 2396

QY 3281 aagagcccgactcactagaagacgtggtcttcaaaagagctcttccagacacccaggtc 3340
Db 2397 AAGAGGCCCAGTCTACATAGAGACCTTGGCTGGCTTCAAAAGAGCTCTTCCAGACACACAGGTC 2456
```

QY	3341	ctcttgaggaa	tcaatgactgatgagaaactaccaaaatagcctgcaaatctccaccac	3400
Db	2457	CCTCTGAGGAATCAATGACTGATGAGAAAATACCAAAATAGCCTGCAAAATCTCCACCAC	2516	
QY	3401	cagaatcaatgagacactccaacagcacaaaagcaaatggcctaagagaagtctcaggaaaag	3460	
Db	2517	CAGAATCAGTGAGACACTCCAACAGCACAAAAGCAATGGGCTTAAGGAAGTCTCAGGAAG	2576	
QY	3461	cagatgtagagaaatattcttagcactcaggaaactaaacacatcagcagggaagcca	3520	
Db	2577	CAGATGTAGAGAGAAATTTCTTAGCACTCAGGAATTAACACATCAGCAGGGAAGCCA	2636	
QY	3521	tgcttacgcccacacagcaggagtgatgagaaagacattaaagcatttatgggaactc	3580	
Db	2637	TGCTTACGCCCAACACAGCAGGAGGTGATGAGAAAAGACATTAAGCATTTATGGGAATC	2696	
QY	3581	caatgcagaactggaactggcaggaaactttaccctggcagcaagaagacagctacagactc	3640	
Db	2697	CAGTCAGAAAATCGGACCTGGCAGGAACCTTTACCTGGCAGCAAAAGACAGCTTACAGACTC	2756	
QY	3641	ctaaggaaaagccagggctctagaagacctggctggctttaagagactcttccagactc	3700	
Db	2757	CTAAGGAAAAGCCACGAGCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACTC	2816	
QY	3701	ctggctcaccagggaattagtggtctggtggttaaacaccactaaaatcacctcgactctc	3760	
Db	2817	CTGGTCACACCCAGGAATTTAGTGGCTGTGTTGTTAAACCACTAAAATACCTGCGACTCTC	2876	
QY	3761	cacagtcaagccagtggacaccccaacagcacaaagcaagcaagcaccagcaagaagtatca	3820	
Db	2877	CACAGTCACAGCCAGTGGACACCCCAACAGCACAAAGCAACGACCCCAAGAGAAATGATCA	2936	
QY	3821	ggaagcagatgtagaggagaactcttagcgtgcaggaaacttaatggccatcagcaggca	3880	
Db	2937	GGAAAGCAGATGTAGAGGGAGAACTTATAGCGTGCAGGAATCTAATGCCATCAGCAGGCA	2996	
QY	3881	aagccatgcacacgcctaaaccatcagtaggtggaagagaaaagacatcatatttggg	3940	
Db	2997	AAGCCATGCACACGGCTAAACCATCAGTAGGTGAAGAGAAAGACATCATATTTGTGG	3056	
QY	3941	gaactccagtgcagaaactggacctgacagagaacttaaccggcagcaagagacggccac	4000	
Db	3057	GAACTCCAGTCAGAAAATCGGACCTTGACAGAGAACTTAACCGGCAGCAGAGACGGCCAC	3116	
QY	4001	aaactcttaaggaagggcccaggctctggaagacctgactggctttaagagctctcc	4060	
Db	3117	AAACTCCTAAGGAAGGCCAGGCTCTGGAAGACCTGACTGGCTTTAAGAGCTCTTCC	3176	
QY	4061	agacccctggctactgaagcagtggtgctgctggcaaaaactaactaaaatgcctgcg	4120	
Db	3177	AGACCCCTGGTCATACTGAAGAAGCAGTGGCTGCTGGCAAAACTACTAAAATGCCCTGG	3236	
QY	4121	aatcttctccagaaatcagcagacaccccacaaagcacagaagaaggcagcccaagacac	4180	
Db	3237	AATCTTCTCCACAGAATCAGCAGACACCCCAACAGCACAGAAGGCGAGCCCAAGACAC	3296	
QY	4181	ctttggaaaaagggacgtcacagaagagctctcagccctgaagagctcacacagacat	4240	
Db	3297	CTTTGGAGAAAAGGACGTACAGAGGAGCTCTCAGCCCTGAAGAGCTCACACAGACAT	3356	
QY	4241	caggggaaccacacacagataaagtaccaggaggtgagataaaaagcatcaacgcgt	4300	
Db	3357	CAGGGAAAACCCACACACAGATAAAGTACCAGGAGGTGAGGATAAAAGCATCAACGCCGT	3416	
QY	4301	ttagggaactgcaaaacagaacttgaccacagcagcaagtgttaactggtagcaagagcg	4360	
Db	3417	TTAGGGAAAATGCAAAACAGAACTTGACCCAGCAGCAAGTGTAACTGGTAGCAAGAGCG	3476	
QY	4361	accctaaaactaaggaagggcccaccccttagaagacctggtggtggaagagctct	4420	
Db	3477	ACCCAAAANACTAAGGAAAAGGCCAACCCCTAGAAAGACCTGGCTGGCTGGAAGAGCTCT	3536	
QY	4421	tcagacaccagtatgctacgaagcccacgactcagcagaaaaactaccaaaatagcct	4480	

Db	3537	TCAGACACCAAGTATGACTGACAAGCCACGACTCACGAGAAAACCTACCAAAATAGCCT	3596
QY	4481	gcagatcaacaacagacccagtgacacaccaacaagctccaagccacagtccaagagaa	4540
Db	3597	GCAGATCAACAACAGACCCAGTGGACACACCAACAAGCTCCAAGCCACAGTCCAAGAGAA	3656
QY	4541	gtctcaggaaaagtggcgtlagaagaagattcttcgcactcaggaaacgaacaccactcag	4600
Db	3657	GTCTCAGGAAAGTGGAGCTTAGAAGAGAATTTCTTCGCACCTCAGGAACGAACACCATCAG	3716
QY	4601	caggcaaaagccatgcacacaccccaaccaggcagtaagtgtgagaaaaacatctacgcct	4660
Db	3717	CAGGCAAAAGCCATGCAACACACCCCAACCCAGCTTAAGTGTGAGAAAAACATCTACGCAT	3776
QY	4661	ttatgggaactccagtgcagaaactggaactgacagagaacttaactggcagcaagac	4720
Db	3777	TTATGGGAATCTCCAGTGCAGAAACTGGACCTGACAGAAACTTAACTGGCAGCAAGAGAC	3836
QY	4721	ggctcaaaactcctaaggaaaagcccaggctctagaagacctggctggctttaaagac	4780
Db	3837	GGCTACAAACTCCTAAGGAAAAGGCCAGGCTCTAGAAGACCTGGCTGGCTTTAAAGAGC	3896
QY	4781	tcttcacagacagaggtcacactgaggaaatcaatgactaacgataaaactgccaaagttag	4840
Db	3897	TCATTCCAGACACACGAGGTCACTGAGGAATCAATGACTTAACGATAAACTGCCAAAGTAG	3956
QY	4841	cttgcaaatcttcacaaccagacctagacaaaaccagcaagctccaagcagcgctca	4900
Db	3957	CTTGCAAAATCTTCACAAACAGACACTAGACAAAACCCAGCAAGCTCCAAGCGAGGCTCA	4016
QY	4901	agacatccctggggaagtggcggtgaagaagagctccctagcagttggcaagctcacac	4960
Db	4017	AGACATCCCTGGGAAAGTGGCGGTGAAGAAGAGCTCCTAGCAGTTGGCAAGCTCACAC	4076
QY	4961	agacatcaggagagactacacacacacacacagagcgaacagagagatggttaagagcatga	5020
Db	4077	AGACATCAGGAGAGACTACACACACACACAGAGGCAACAGAGAGATGGTAAGAGCATGA	4136
QY	5021	aagcatttatgaggttccaagaagcagacttagactcagcagcaagtcttaactggcagca	5080
Db	4137	AGCATTTATGGAGTCTCCAAGACGACTTTAGACTCAGCAGCAAGTCTAACTTGGCAGCA	4196
QY	5081	agaggcagctgagaactccttaagggaagctgaagtccctgaagacctggcgcgcttca	5140
Db	4197	AGAGGCAGCTGAGAACTCCTAAGGGAAGTCTGAGTCCCTGAAGACCTGGCCGGCTTCA	4256
QY	5141	tcgagctctccagacaccaagttcacactaaggaaatcaatgactaatgaaaaaactacca	5200
Db	4257	TCAGGCTCTCCAGACACCAAGTCACTAAGGAATCAATGACTTAATGAAAAAACTACCA	4316
QY	5201	aagtatcctcacagacttcacagccagacctagtcagaccccccaacaagctccaagccac	5260
Db	4317	AGTATCTCCACAGAGTCTACAGCCAGACCTAGTGGACACCCCAACAAAGTCCAAGCCAC	4376
QY	5261	agccccagagaagtctcaggaagcagacactgaagaagaatttttagcatttaggaaac	5320
Db	4377	AGCCCAAGAGAACTCTCAGGAAGCAGACACTCAGAAGAATTTTAGCATTTAGGAAAC	4436
QY	5321	aaacgcctaagcaggcaaaagccatgcacacaccccccaaccagcagtaggtggaagagaag	5380
Db	4437	AAACGCATCAGCAGGCAAAAGCCATGCACACACCCCAACCCAGCAGTAGGTGGAAGAGAAAG	4496
QY	5381	acatcaacacgtttttgggaactccagtcagaaactggaccagccaggaatttaacctg	5440
Db	4497	ACATCAACACGTTTTTTGGGAACCTCCAGTGCAGAAACTTGGACCCAGGAAATTTTAACTG	4556
QY	5441	gcagcaatagacgcgtcacaaactcgttaaggaaaagcccaggctctagaagaactgactg	5500
Db	4557	GCAGCAATAGACGGCTTACAAACTCTGTAAGGAAAAGGCCAGGCTCTAGAGAAGACTGACTG	4616
QY	5501	gcttcagagagcttttccagacaccatgcactgataacccccacagctgtatgagaaaacta	5560

.Db 4617 GCTTCAGAGAGCTTTTCCAGACACCATTGCAGTATCAACCCACAGCTGATGAGAAACTA 4676
QY 5561 ccaaaaaa tactgtcaaatctccycaatcagacccagcgagaccccccaacacaaa 5620
Db 4677 CCAAAAAAATAC TCTGCAAAATCTCCGCAATCAGACCCAGCGACACCCCAACAAACACAA 4736
QY 5621 agcaacggcccaagagaagcctcaagaaaaagacagcgttagaggagaatttttagcattca 5680
Db 4737 AGCAACGGCCCAAGAGAAAGCCTCAAGAAAGCAGACGTAGAGGAAGAAATTTTATGCAATCA 4796
QY 5681 ggaactaaacacatcatcagcaggcaaaagccatgcacacgcctaaagcagcagtagtgaaag 5740
Db 4797 GGAACCTAACCCATCAGCAGGCAAAAGCCATGCACAGCCCTAAAGCAGCAGTAGAGTGAAG 4856
QY 5741 agaagacatcaacacatttgggggactccagtggaagaaactggagaaactgcttaggaatt 5800
Db 4857 AGAAGACATCAACACATTTTGGGGACTCCAGTGGAGAAACTGGACCTCTAGGAATTT 4916
QY 5801 taactggcagcagagcggcccaaaactcctctaaagaaaaaggcgaaggtctctagaagatc 5860
Db 4917 TACTGGCAGCAAGAGACGGCCCAAAACTCTTAAGAAAAAGGCCAAGGCCTCTAGAAGATC 4976
QY 5861 tggctggcttcaagagctcttcagacacccaggtcacactgaggaatcaatgacagatg 5920
Db 4977 TGGCTGGCTTCAAAGAGCTCTTCAGACACCCAGTCACTAGGAATCAATGACCGATG 5036
QY 5921 acaaaatcacagaagtatcctgtcaaaatctccacaacccagacccagtcacaaaccccaacaa 5980
Db 5037 ACAAAATCACAGAAGTATCTCGCAAAATCTCCACAACCCAGACCCAGTCAAAACCCCAACAA 5096
QY 5981 gctccaagcaactcaagatataccttggggaaagttaggtgtgaaagaagaggtctctac 6040
Db 5097 GCTCCAAGCAACGACTCAAGATATCTCTTGGGGAAAGTAGTGTGTCAAAAGAAAGGTCCTAC 5156
QY 6041 cagtccggaactcacacagcgtccagggaagaccacacacagacacacagagagacagcag 6100
Db 5157 CAGTCGCAGACTCACACACAGCTCAGGGAAGACCACACAGACACAGACAGACAGCAG 5216
QY 6101 gagatgaaagagcatcaaaagcgtttaaggaattctgcaagcagatgctggacccagcaa 6160
Db 5217 GAGATGAAAGAGCATCAAAAGCGTTTAAAGGAATCTGCAAAAGCAGATGCTGGACCCAGCAA 5276
QY 6161 actatgaaactggatggagaggtggccaaagaacacctaaaggagagggcccaatcactag 6220
Db 5277 ACTATGGAATGGGATGGAGGTGGCCAAAGAACACCTTAAGGAAGAGGGCCCAATCACTAG 5336
QY 6221 aagacctggcggcttcaagagctcttcagacacacagacacacactgaggaatcaacaa 6280
Db 5337 AAGACCTGGCGGCTTCAAGAGACTCTTCAGACACCCAGACACACTGAGGAATCAACAA 5396
QY 6281 ctgatgacaaaactaccaaaatagcctgcaaaatctccaccaccagaatcaatggacactc 6340
Db 5397 CTGATGACAAAACCTACCAAAATAGCCTGCAAAATCTCCACCACCAAGATCAATGGACACTC 5456
QY 6341 caacaagcaacaggcggcccaaaacaccttggggaaaaaggatatagttgaagagc 6400
Db 5457 CAACAAGCAACAGGAGCGGCCCAAAACACCTTTGGGGAAAAGGGATATATAGTGGAAAGC 5516
QY 6401 tctcagccctgaagcagctcacacagaccacacacacagacaaagtaccagggagatgagg 6460
Db 5517 TCTCAGCCCTGAAGCAGCTCACAGACACCACACACAGACAAAGTATACCAGGAGATGAGG 5576
QY 6461 ataaggcatcaacgtgttcagggaactgcataaacagaaactggacccagcagcgaagt 6520
Db 5577 ATAAGGCATCAACGTTTCAGGGAAACTGCAAAACAGAAACTGGACCCAGCAGCAAGTG 5636
QY 6521 taactggtagcaagaggcagcaagaactcctaaaggaaaaagcccaacccctagaagact 6580
Db 5637 TAACTGGTAGCAAGAGCGACCAAGAACTCCTTAAGGGAAAAGGCCAACCCCTTAGAAGACT 5696
QY 6581 tggctggcttgaagagctcttcagacacccagtatagtcactgacaagccacagactcacg 6640
Db 5697 TGGCTGGCTTGAAGAGACTTTCAGACACACCATGATATGACTTGACACAGCCCACTACG 5756

QY 6641 agaaaaactacaaaatagcctgcagatctccacaaccagaccccgtaggggtacccccaaaa 6700
Db 5757 AGAAAAC TACCAAAATAGCCTGCAGATCTCCACAACCCAGACCCAGTGGGTACCCCAACAA 5816
QY 6701 tcttcaagccacagtcacaagagaagtctcaggaaaagcagacgttagagaagaagaatccttag 6760
Db 5817 TCTTCAAGCCACAGTCCCAAGAGAAGTCTCAGGAAAGCAGCGTAGAGGAAGAATTCCTTAG 5876
QY 6761 cactcaggaaacgaacaccatcagtagggaagctaaggacacaccccaaacacagcagag 6820
Db 5877 CACTCAGGAACGACACCATCAGTAGGGAAGCTATGGACACACCCCAACCCAGCAGAG 5936
QY 6821 gtgatgagaagacatgaaagcatttatgggaactcctcagtgcaaaatttggaacctgcccag 6880
Db 5937 GTGATGAGAAGACATGAAAGCATTTATGGGAAC TCCAGTCCAGTTGCAGAAATTTGACCTGCCAG 5996
QY 6881 gaaatttacctgycagcaaaagatggccacaacactcctaaaggaaaaagcccgactctag 6940
Db 5997 GAAATTTACCTGGCAGCAAAAGATGGCCACAAACTCCTTAAGGAAAAGGCCCGCTCTAG 6056
QY 6941 aagacctggctggcttcaagagctcttcagacacccagcactgacaagcccaacagctg 7000
Db 6057 AAGACCTGGCTGGCTTCAAGAGCTCTTCAGACACCCAGCAGCTGACAGACCCACGACTG 6116
QY 7001 atgagaaaaactacaaaatagcctgcaaaatctccacaacccagaccccgtagcaccccg 7060
Db 6117 ATGAGAAAAC TACCAAAATAGCCTGCAAAATCTCCACAACCCAGACCCAGTGGACACCCCGAG 6176
QY 7061 caagcaacaagcaacggcccaagagaacactcagggaaaagcagacgttagagaagaataattt 7120
Db 6177 CAAGCAAAAGCAACGGCCCAAGAGAAACCTCAGGAAAAGCAGACGTAGAGAGAAGAAATTTT 6236
QY 7121 tagcactcaggaacgaacaccatcagcaggcaaaagccaaggacaccccaaaaacacagcag 7180
Db 6237 TAGCACTCAGGAACGAAACACCATCAGCAGGCAAAAGCCATGGACACCCCAACCCAGCAG 6296
QY 7181 taagtgtatgagaaaaatatcaacacatttgtgaaaactccagtgcaaaactggacactgac 7240
Db 6297 TAAAGTGTATGAGAAAATATCAACACATTTTGTGGAACATCCAGTGCAGAAAAC TGGACCTGC 6356
QY 7241 taggaatttacctgycagcaagacagacccagactcctaaaggaaaagcctgaagctc 7300
Db 6357 TAGGAATTTACCTTGGCAGCAAGAGACAGCCACAGACTCTCTTAAGGAAAAGCTGAGGCTC 6416
QY 7301 tagaggacctgtgtgcttcaagaaactctccagacacccaggtcacaactgaggaatacaa 7360
Db 6417 TAGAGGACCTGGTTGGCTTCAAGAACTCTTCAGACACCCAGGTCACACTGAGGAATCAA 6476
QY 7361 tgaactgtacaaaaatcacagaagtaacctgtaaaatctccacagccagatcattcaaaa 7420
Db 6477 TGACTGTATGACAAAATCACAGAAATCTCTGTAAATCTCCACAGCCAGAGTCAATTCAAAA 6536
QY 7421 cctcaagaaactccaagcaaaagctcaagataccccctgttgaaaagtggacatgaaaagaag 7480
Db 6537 CCTCAAGAGACTCCAAGCAAAAGGCTCAAGATACCCCTTGGTGAAGTGGACATGAAGAAG 6596
QY 7481 agcccttagcagtcagcaagctcacacgggacatacaggggagactcacgcaaacacacacag 7540
Db 6597 AGCCCTTAGCAGTCAAGCAAGCTCACAGGACATCAGGGGAGACTACGCCAAACACACACAG 6656
QY 7541 agccaacagagatagtagagcatcaaaagcgtttaaggaggtctccaaaagcagatccttg 7600
Db 6657 AGCCCAACAGGAGATAGTAAAGAGCATCAAAAGCGTTTAAAGGAGTCTCCAAAGCAGATCCTCG 6716
QY 7601 acccagcagcaagtgaactggtagcaggagcgactgagaactctgaagaaaaagccc 7660
Db 6717 ACCCAGCAGCAAGTGTAACTGGTAGCAGGAGGAGCTGTGAGAACTCTGTAAAGAAAAGGCC 6776
QY 7661 gtgctctagaagacctggttgacttcaaaagactctctcctcagcaccaggtcacaactgaag 7720
Db 6777 GTGCTCTAGAAGACCTGGTTGACTTCAAGAGAGCTCTTCTCAGCACCCAGGTCACACTCAAG 6896

QY	7721	agtcaatgactattgacaaaaacacaaaaatccctcgcaattctccccaccagaaactaa	7780
Db	6837	AGTCATGACTATTGACAAAAACAAAAAATTCCTTGCAATCTCCCCACCAGAACTAA	6896
QY	7781	caqacactccacgagcacaaagatgcccaaaacacgtccccagaaagaagtaaaag	7840
Db	6897	CAGACTGCCACGAGCACAAAGATGCCCAAGACAGCTCCAGGAAGAAGTAAAG	6956
QY	7841	aggagctctcagcagttgagggctcacgcaaacatcagggcaaaagcacacacacaca	7900
Db	6957	AGGAGCTCTCAGCAGTTGAGGGCTCACGCCAAACATCAGGGCAAAAGCACACACACA	7016
QY	7901	aggaaccagcaagcggtgatgsggcatacaagttatgaagcaacgtgcaaaagaagaac	7960
Db	7017	AAGAACCACGACGCGTGATGAGGCATCAAAAGTATTGAAGCAACGTGCAAAAGAGAAAC	7076
QY	7961	caaacccagtagaaggaaccacgacgagagaagcccaagcagcacctaaggaagccccc	8020
Db	7077	CAAACCCAGTAGAAGAGGAACCCAGCAGGAGGAAGGCCAAGACACTAAGGAAAAGGCC	7136
QY	8021	aaccctggaaagacctggccggtctcacagagctctctgaaacatacaggtcacactcag	8080
Db	7137	AACCCCTGGAAGACCTGGCCGCTTCACAGAGCTCTCTGAAACATCAGGTACACTCAGG	7196
QY	8081	aatcactgactctggcaagccactaaataaccctcgcaattctccccacttagaagtgg	8140
Db	7197	AATCATGACTGCTGGCAAAAGCCACTAAATATACCTTCGCAATCTCCCCACTAGAAATGG	7256
QY	8141	tagacacacagacagacaaagagcctctcagacacgtgtgcagaaggtacaagtaa	8200
Db	7257	TAGACACCACAGACACAAAGAGGCATCTCAGGACAGCTGTGCAAGAGGTACAAAGTAA	7316
QY	8201	aagaagagctctcagcagtcgaagttcacacaaacatcaggggaaacacaggtgcagaca	8260
Db	7317	AAGAAGAGCTTCACAGCTCAAGTTACACACAACATCAGGGGAACCCAGGATCGACACA	7376
QY	8261	agaaaccagcaggtgaagataaaggcatcaaaagcattgaaggaattcgcaaacacagcac	8320
Db	7377	AAGAACCCAGCAGGTGAAGATAAAGGCATCAAAAGCATTTGAAGGAATCTGCAAAACAGACAC	7436
QY	8321	cggtccacgacgaagtgaactgcagcagagagcgcgaagacacccagggaaagtg	8380
Db	7437	CGGCTCCACGACGAGTGAATTCGCGAGGAGAGCGGCCAAGACCCAGGGAAAGTG	7496
QY	8381	cccaagccataagaacacctagctggtctcaaaagaccacgacgaggtcacactgaagaat	8440
Db	7497	CCCAAGCCATAGAAGACCTAGCTGGCTTCAAGACCCACGACGAGTACACTGAAGANT	7556
QY	8441	caatgactratgacaaaaaccataataaccctgcaaatcaaccacagactagaagaca	8500
Db	7557	CAATGACTGATGACAAAAACCACTAAATATACCTTGCAAAATCATCCACGAACCTAGAAAGACA	7616
QY	8501	ccgcaacaagctcaagagacgcccagacacgtgcccagaaagttagaagtgaaggagg	8560
Db	7617	CCGCAACAAGCTCAAGAGACGGCCAGACACGTGCCCCAGAAAGTAGAAGTGAAGGAGG	7676
QY	8561	agctgttagcagttggcgaagctcacacaaacctcaggggagacacgcacaccgacaaaq	8620
Db	7677	AGCTGTATACAGTTCGCAAGCTCACACAACCTCAGGGGAGACCAGCACACCCACAAG	7736
QY	8621	agccggtaggtgaagggcaagcagaaagcatttaagcaacctgcaaaagcggaacgtgg	8680
Db	7737	AGCCGGTAGGTGAGGGCAAAAGGCACAAAAGCATTTAAGCAACGTGCAAAAGCGAACGTGG	7796
QY	8681	acqgaagaatgataatggcacagagacagcacaagcacctaaggaagaaagggcccaac	8740
Db	7797	ACGCAAGAATGTAAATTGGCAGCAGGAGACGCCAAGACACCTAAGGAAAAGGCCCAAC	7856
QY	8741	ccctggaaacacctggccagcttccaaagagctctctcaaacacacagggccacactgaggaac	8800
Db	7857	CCCTGGAAGACCTGGCCACTTCCAGAGCTCTCTCAAAACACAGGCCACACTGAGGAAC	7916
QY	8801	tggcaaatggtgctgtagctttacaagcgctcccaagcacaacacacctgacagtgga	8860

Db	7917	TGCAAAATGGTGTGCTGATAGCTTTACAAAGCGCTCCAAAAGCAAAACACCTTGACAGTGGAA	7976
QY	8861	aaccttaaaaaatatcagaagagttcttcggccctaaagtagaacccgttggaacg	8920
Db	7977	AACCTCTAAAAATATCCAGAAGAGTTCTTCGGGCCCTTAAGGTAGAACCCCGTGGGACCG	8036
QY	8921	tgtgaagcaccagagacccctgtaaaaatcacaagcaaaaagcaaaccttccctgccccac	8980
Db	8037	TGTTAAGCACACAGAGACCCCTGTAAAATACAAAAGCAAAACAACTTCCCTGCCCCAC	8096
QY	8981	tgccttcaagaggagggtggaagatggaagagtcacgggaacccaagagcgtcgct	9040
Db	8097	TGCCCCCTCAAGAGGGAGGTGGCAAAAGATGGAACGCTCACGGGAACCAAGAGCGTGCCT	8156
QY	9041	gcataccagcaccagaggaattgtgagagagctgcacgccaagaaagacagaggttg	9100
Db	8157	GCATGCCACACCAGAGGAATTTGTGGAGGAGCTGCCAGCCACGAAGACAGAGGGTTG	8216
QY	9101	ctccaggggcaagaggaataatcatccgaacccgtggtcatcatgaagagagtttgagga	9160
Db	8217	CTCCCAGGCAAGAGGCAAAATCATCCGAACCCGTGGTTCATCATGAAGAGAAGTTTGAGGA	8276
QY	9161	cttctgcaaaaagaattgaacctgcggaagagctgaacacgacacatgaacccaaca	9220
Db	8277	CTTCTGCAAAAAGAAATTTGAACCTTGCGAAGAGCTGAACAGACGACATGAAAACCAACA	8336
QY	9221	agagaaacacaaattacaagactcggtccctgaaataaagggaatatccctcgctcca	9280
Db	8337	AAGAGAAACAAATTTACAAGACTCGGTCCCTGAAAATAAGGAATATCCCTCGCTCCA	8396
QY	9281	gacgcaagataagactgagcgagacagcaataaactaggtctttgtattagcaaaa	9340
Db	8397	GAGCCCAAGATAAGACTGAGCGAGACAGCAATAAATACTGAGGTCTTTGTATTAGCAAAA	8456
QY	9341	gaatgaataaaccagaaattgaagaagcccatgaagacctccccagagatggacattc	9400
Db	8457	GAATAGAAATAAACAGAAAATGAAAAGAAAGCCCATGAAGACCTCCCCAGAGATGGACATTC	8516
QY	9401	agaatccagatgatgagcccgaaccctacactagacaaagtctactgagaaacaaa	9460
Db	8517	AGAAATCCAGATGATGGAGCCCGGAACCCATACCTAGAGACAAGTCACTTGAGAACAAA	8576
QY	9461	ggtgcttgaggtctgctagacagaatgagagctcccagctcctgaagtcggcagagagagcg	9520
Db	8577	GGTGCTTGAGGTCTGTAGACAGAAATGAGAGCTCCAGCCCTAAGGTGCGCAGAGAGCG	8636
QY	9521	gagggcagaagagtcggaaggttctcatgcagaatacgaagaagggaagggaagcagaa	9580
Db	8637	GAGGGCAGAAGAGTGCGAAGGTTCTCATGCAAGATCAGAAGGGAAAGAGGACAGGAA	8696
QY	9581	attcacactlccatgctcctgagatcaagaagaacaaaagccagcctgcagcagcactt	9640
Db	8697	ATTCAGACTCCATGTGCTGAGATCAAGAAAGACAAAAGCCAGCCCTGCAAGACGACTT	8756
QY	9641	tggagcgaataatgtgcagagagtaacgcggaggtgccaagaggtgtgcagaaatccaa	9700
Db	8757	TGGAGAGCAAAATCTGTGCAGAGAGTAACGCGGAGTGTCAAGAGGTGTGCAGAAAATCCAA	8816
QY	9701	agaagctgaggaacaatgtgtgtcaagaaaaatacaacccaagaagtca tagggaacgtg	9760
Db	8817	AGAAGCTGAGGCAATGTGTGTCAAGAAAAATAACACCAAGATCATAGGACAGTG	8876
QY	9761	aaqata 9766	
Db	8877	AGATA 8882	

RESULT 5

HSAK67

LOCUS

DEFINITION

ACCESSION

H. sapiens gene for antigen of the monoclonal antibody Ki-67.
X74107

29965 bp DNA linear PRI 25-JUN-1997

```

VERSION      X74107.1  GI:1370102
KEYWORDS     antigen; monoclonal antibody; nuclear protein; tandem repeat.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 29965)
AUTHORS     Gerdes,J.
TITLE       Direct Submission
JOURNAL     Submitted (12-JUL-1993) J. Gerdes, Forschungsinstitut Borstel, Div.
REMARK      Molecular Immunol., Parkallee 22, 23845 Borstel, FRG
REFERENCE    2 (bases 1 to 29965)
AUTHORS     Duchrow,M., Schluter,C., Wohlenberg,C., Flad,H.D. and Gerdes,J.
TITLE       Molecular characterization of the gene locus of the human cell
            proliferation-associated nuclear protein defined by monoclonal
            antibody Ki-67
JOURNAL     Cell Prolif. 29 (1), 1-12 (1996)
MEDLINE     96177792
FEATURES     Location/Qualifiers
             source      1..29965
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="10"
                        /cell_line="IM9"
                        /map="q25-qter"
             misc_feature 1..729
                        /note="CpG rich island"
             exon         75..180
                        /number=1
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       181..522
                        /number=1
                        /evidence=experimental
             exon         523..703
                        /number=2
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             repeat_unit  545..568
                        /note="box 1"
             repeat_unit  569..592
                        /note="box 2"
             intron       704..3100
                        /number=2
                        /evidence=experimental
             exon         3101..3179
                        /number=3
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       3180..3273
                        /number=3
                        /evidence=experimental
             exon         3274..3389
                        /number=4
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       3390..6958
                        /number=4
                        /evidence=experimental
             exon         6959..7025
                        /number=5
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       7026..9741
                        /number=5
                        /evidence=experimental
             exon         9742..9787
                        /number=6
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       9788..10270
                        /number=6
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             repeat_unit  10271..11350
                        /number=7
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       11351..12673
                        /number=7
                        /evidence=experimental
             exon         12674..12849
                        /number=8
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       12850..13825
                        /number=8
                        /evidence=experimental
             exon         13826..14138
                        /number=9
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       14139..14225
                        /number=9
                        /evidence=experimental
             exon         14226..14344
                        /number=10
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       14345..14454
                        /number=10
                        /evidence=experimental
             exon         14455..14626
                        /number=11
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       14627..15737
                        /number=11
                        /evidence=experimental
             exon         15738..15893
                        /number=12
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             intron       15894..16847
                        /number=12
                        /evidence=experimental
             exon         16848..23692
                        /number=13
                        /evidence=experimental
                        /product="antigen of the monoclonal antibody Ki-67"
             repeat_unit  17429..17794
                        /note="Ki-67 no 1"
             repeat_unit  17795..18159
                        /note="Ki-67 no 2"
             repeat_unit  18160..18586
                        /note="Ki-67 no 3"
             repeat_unit  18587..18889
                        /note="Ki-67 no 4"
             repeat_unit  18890..19252
                        /note="Ki-67 no 5"
             repeat_unit  19253..19618
                        /note="Ki-67 no 6"
             repeat_unit  19619..19981
                        /note="Ki-67 no 7"
             repeat_unit  19982..20349
                        /note="Ki-67 no 8"
             repeat_unit  20350..20717
                        /note="Ki-67 no 9"
             repeat_unit  20718..21070
                        /note="Ki-67 no 10"
             repeat_unit  21071..21433
                        /note="Ki-67 no 11"
             repeat_unit  21434..21799
                        /note="Ki-67 no 12"
             repeat_unit  21800..22165
                        /note="Ki-67 no 13"

```



```
repeat_unit      22166. .22585
repeat_unit      /note="Ki-67 no 14"
repeat_unit      22586. .22882
repeat_unit      /note="Ki-67 no 15"
repeat_unit      22883. .23239
intron           /note="Ki-67 no 16"
intron           23693. .24569
intron           /number=13
exon             /evidence=experimental
exon             24570. .25013
exon             /evidence=experimental
exon             /number=14
exon             /evidence=experimental
exon             /number=15
polyA_signal     /product="antigen of the monoclonal antibody Ki-67"
polyA_signal     27102. .27107
polyA_signal     27637. .27642
polyA_signal     /note="inactive"
polyA_signal     28345. .28350
polyA_signal     29576. .29581
BASE COUNT      8995 a 6446 c 6578 g 7946 t
ORIGIN
```

```
Query Match      70.1%; Score 6847; DB 9; Length 29965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6847; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2416 ggaagaaatgtcttcagtcacagaatgcagcaaacacagccatctgataaatgctct 2475
Db  16847 GGAGGAATGTCTTTCAGTCACAGATGCGCAAAACAGCCATCTGATAAATGCTCT 16906

Qy  2476 gcaagccctcccttaagcggcagtgattagagaaaaatggaacgtgcagaaaaacgccc 2535
Dy  16907 GCAAGCCCTCCCTTAAGACGGCAGTGTATTAGAGAAATGGAACGTAGCAAAACGCCCC 16966

Qy  2536 aggaacacactacaataatgactctctgtgagacaaaacttcagatactagacagagcct 2595
Dy  16967 AGGAACACCTACAAATGACTTCTCTGGAGACAAAACCTTCAGATACTGAGACAGAGCCT 17026

Qy  2596 tcaaaaacagtatccactgtaaacaggtcaggaagttctacagagttcaggaatatcacg 2655
Dy  17027 TCAAAAACAGTATCCACTGTAAACAGGTCAGGAAGGCTTACAGAGTTTCAGGAATATACAG 17086

Qy  2656 aagctacctgtggaagtgaagtaagtaagaaacaaatacagaaattgttgagtgcatccta 2715
Dy  17087 AAGCTACCTGTGGAAGTAAAGAGTGAAGAAACAAATACAGAAATTTGTTGAGTGCATCCTA 17146

Qy  2716 aaaaaggttcagaagcgaacactactactactacaaagagagagaagagagataagaaata 2775
Dy  17147 AAAAGAGGTCAGAGGCAACACTACTACAAACAAAGAGAGAGAGATGAAGGAAATA 17206

Qy  2776 gaaagaccttttgacacataaaggaataattgaattaaaagaaaacgagtaaaaagatg 2835
Dy  17207 GAAAGACCTTTTGACACATATPAGGNAATATTTGAATTAAGAAACACGATGAAAAGATG 17266

Qy  2836 aagcaaatgaagagatcagaacttgggggcagaaatgtgcaccataatgtctgacctgaca 2895
Dy  17267 AAAGCAATGAAGAGATCAAGAACTTTGGGGCAGAAATGTGCACCAATGTCTGACCTGACA 17326

Qy  2896 gacctcaagagcttgcctgtatatacagaactcagaaagacacagcgcgtggccagaaatctc 2955
Dy  17327 GACCTCAAGAGCTTCCTCGATACAGAACTCATCAAGAGACACGGCAGCTGGCCAGAAATCTC 17386

Qy  2956 ctccaaacccagatcatgccaagcaccacaaagagtgagaaagggcaaaatcactaaaatg 3015
Dy  17387 CTCCAAACCAAGATCAGCCCAAGGCCAACAAAGATGAGAAAGGGCAAAATCACTAAAATG 17446
```

```
Qy  3016 cctgtccagtcattacaaccagaaccaataaaacaccccccaacacacacacacacagttg 3075
Dy  17447 CCTGCGAGTCATTACAACAGAACCAATAAACACCCCAACACACACACAAACAGTTG 17506

Qy  3076 aaggcatcctcctgggaaagtaggtgaaagaagaagcctcctagcagtcaggcaagttcaca 3135
Dy  17507 AAGGCATCCTCGGGAAGTAGGTGTGAAGAAGAGAGCTCCTAGCAGTCGGCAAGTTTACA 17566

Qy  3136 cggacgtcagggagagaccacgcacacgcacagagagagccagcagcagagatggcaagagcatc 3195
Dy  17567 CGGACGTCAGGGAGAGACACGACACAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 17626

Qy  3196 agaacgtttaaggagtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3255
Dy  17627 AGAAGCTTTAAGGAGTCTCCAAAGCAGATCCTGGACCCAGCAGCCCGTGTAACTGGAATG 17686

Qy  3256 aagaagtggcccaagaagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3315
Dy  17687 AAGAAGTGCCCAAGAACGCTTAAGGAAGAGAGGCCAGTCACTAGAAAGACCTGGCTGGCTTC 17746

Qy  3316 aaagagctcttcacagacacacaggtcctcctcctcctcctcctcctcctcctcctcctcctc 3375
Dy  17747 AAAGAGCTCTTCCAGACACACAGGTCCCTCTGAGGAATCAATGACTGATGAGAAACTACC 17806

Qy  3376 aaaaatagcctgcacaaatctccacacacagaaatcagtgagacactcccaacagcacaagcaaa 3435
Dy  17807 AAAATAGCCTGCAAAATCTCCACCACCAGAAATCAGTGGACACTCCAACAAGCACAAAGCAA 17866

Qy  3436 tggcctaagagaagatctcaggaagaaagcagatgtagagaaagaatttcttagcactcaggaaa 3495
Dy  17867 TGGCCTTAAGAGAAGTCTTCAGGAAGACAGATGTAGAGAAAGAAATTTCTTAGCACTCAGGAAA 17926

Qy  3496 ctacaacacatcagcagggaaagccatgcttcagcgcacacccagcagcaggtgatgagaaa 3555
Dy  17927 CTACACCATCATCAGCAGGGAAGCCATGCTTACGCCCAAAACACAGCAGAGGTGATGAGAAA 17986

Qy  3556 gacataaagcatttatgggaaactccagtcgagaaactggacctggcaggaactttacct 3615
Dy  17987 GACATTAAGACATTTATGGGAATCTCCAGTGCAGAAACTGGACCTGGCAGGAACCTTTACCT 18046

Qy  3616 ggcagcaaaagacagctacagactcctcctcctcctcctcctcctcctcctcctcctcctcct 3675
Dy  18047 GGCAGCAAAAGACAGCTACAGACTCCTTAAGGAAGAGGCCAGGCTCTAGAAAGACCTTGCT 18106

Qy  3676 ggccttaagagagctcttcacagctcctggtcacacccagagaaattagtgctgctgtaaa 3735
Dy  18107 GGCCTTAAGAGAGCTCTTCCAGACTCCTGCTCACACCGCAGGAATTAGTGGCTGCTGTA 18166

Qy  3736 accactaaaataccctgcgactctccacagtcagacccagtcagacccacccacacacacacac 3795
Dy  18167 ACCACTAAAATACCCTGCGACTCTCCACAGTCAGACCCAGTCAGACACCCCAACAGACACA 18226

Qy  3796 aagcaacgacccaagaagaagatcaggaagaagcagatgtagaggagaaactcttagcgtgc 3855
Dy  18227 AAGCAACGACCCCAAGAAAGATATCAGGAAGCAGATGTAGAGGGGAACCTCTTAGCGTGC 18286

Qy  3856 aggaatcttaatgccatcagcagcgaagccatgcacacccctaaacccatcagtagtgtaa 3915
Dy  18287 AGGAATCTAATGCCATCAGCAGCAAGCCATGCACACCGCTTAACCATCATCAGTAGTGAA 18346

Qy  3916 gagaagacatcatcatatttgggaaactccagtcagaaactggcagaaactggacctgacagagac 3975
Dy  18347 GAGAAAGACATCATCATATTGTTGGGAATCTCCAGTCAGAAACTGGACCTGACAGAGAAC 18406

Qy  3976 ttaacggcagcaagaagacgcccacaaactcctaaaggaagagcccgagcgtctcggagac 4035
Dy  18407 TTAACGGCAGCAAGAGACGGCCACAACTCCTAAGGAAGAGGCCAGGCTCTTGGGAAGAC 18466

Qy  4036 ctgactggcttttaaagagctcttccagacccctcctggtctactactgaagaagcagtgctgct 4095
Dy  18467 CTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTGTACTACTGAAAGACGATGGCTGCT 18526

Qy  4096 ggcaaaactactaaaatgcctgcgaattcttctccaccagaatcagcagacacaccccaaca 4155
```

Db 18527 GGCAAAACTACTAAATGCCCTGGAAATCTTCCACAGAATCAGCAGACACCCCAACA 18586
QY agcaagaagcgccgaagacacacctttggagaaaggagcgctacagaaggagctctca 4215
Db 18587 AGCACAAGAAGGCAAGCCCAAGACACCTTTGGAGAAAGGAGCGCTACAGAAGGAGCTCTCA 18646
QY gccctgaagaagctcacacagacatcaggggaaacccacacacagataaagtaccagga 4275
Db 18647 GCCCTGAAGAAGCTCACACAGACATCAGGCGAAACCCACACACAGATAAAGTACCAGGA 18706
QY ggtagagataaaagcatcaacgcgttttagggaaaactcaaaaaacagaactggaccacga 4335
Db 18707 GGTGAGGATAAAGCATCAACGGCGTTTATGGGAAACTCCAAACAGAAACTTGGACCCAGCA 18766
QY gcaagttaactggttagcaagagcgaccccaaaactaagaaaaagcccaacccctagaa 4395
Db 18767 GCAAGTGAATGGTAGCAGAGAGCCACCCAAAACAACTAAGGAAAGGCCCAACCCCTAGAA 18826
QY gacctggctggctggaaagagctcttccagacacccagtatgcactgacaagccacgact 4455
Db 18827 GACCTGGCTGGCTGGAAGAGCTTCTCCAGACACACAGTATGCATGACAAGCCACGACT 18886
QY cagagaaaaactaccaaaaatagcctgcagatcacaaaccagaccctggcacaccaaa 4515
Db 18887 CACGAGAAAACTACCAAAATAGCGCTGCAGATCACAAACAGACCCAGTGGACACCAACA 18946
QY agctccaagccacagctccaagagaagctctcaggaagtgtagcgttagagaagtaattctc 4575
Db 18947 AGCTCCAAGCCACAGTCCAAGAGAGTCTCAGGNAAGTGGACGCTAGGAAGAAATTCCTC 19006
QY gcactcaggaaaacacacacacacacacagcggaagcgaatgcacacaccccaaccagagta 4635
Db 19007 GCATCTAGGAAACGAACACCATCAGCAGGCAAGCCATGCACACACCCAAACCCAGCAGTA 19066
QY agtggtagaaaaacatctcgcgatttatgggaactccagtgcagaactggacctgaca 4695
Db 19067 AGTGGTGAGAAAAACAATCTACGGCATTTATGGGAATCTCCAGTGCAGAAACTGGACCTGACA 19126
QY gagaacttaactggcagcaagagacggtctacaactcctcaggaagggcccggtccta 4755
Db 19127 GAGAACTTAATGGCAGCAAGACGGCTACAAACTCTTAAGGAAAGGCCAGGCTCTA 19186
QY gaagacctggctggctttaagagctcttccagacacgaggtcacactcagggaatcaatg 4815
Db 19187 GAAGACCTGGCTGGCTTTAAAGAGCTTCTCCAGACACGAGGTACACTGAGGAATCAATG 19246
QY actaagcataaaactgccaaagtagcctgcaaatcttcacaacagacctagacaaaaac 4875
Db 19247 ACTAACGATTAATGCTCAAAAGTAGCCTGCANAATCTTCAACCCAGACCTTAGACAAAAAC 19306
QY ccagcaagctccaagcgacggctcaagacatccctgggaaagtggcggtgaagaagag 4935
Db 19307 CCAGCAAGCTCCAAGCGACGGCTCAAGACATCCTGGGGAAGTGGCGGTGAAGAAGAG 19366
QY ctctcagcagctggcgaagctcaacagacatcagggagagactacacacacacacagag 4995
Db 19367 CTCTTAGCAGTGGCAAGCTCACACAGACATCAGGAGAGACTACACACACACACACAGAG 19426
QY ccaacgagagatggttaagagcatgaagcatttatggagtctccaaagcagatcttagac 5055
Db 19427 CCAACAGGAGATGGTAGACGCTGAAGCATTTATGGAGTCTCCAAAGCAGATCTTAGAC 19486
QY tcagcgaagcttaactggcagcaagcgagcgactgagaactccttaagggaagctctgaa 5115
Db 19487 TCAGCAGCAAGTCTAACTGCGCAGCAAGAGCGAGCTGAGAACTCCTTAAGGAAAGTCTGAA 19546
QY gtccctgaagacacctgcccggtctcaatcgactcttccagacacaaagtcacataaggaa 5175
Db 19547 GTCCCTGAAGACCTGGCGCGCTTTCATCGAGCTTCTCCAGACACCAAGTCAACATAAGGAA 19606
QY tcaatgacttaataaaaaactaccaagtatcctcagagcttcacagcagacctagtg 5235
.

Db 19607 TCAATGACTAATGAAAAAATACTACAAAGTATCTCTACAGAGCTTCACAGCCAGACTAGTG 19666
QY gacaccccaacaagctcctaagccacagcccaagagagagctctcaggaagcagacactgaa 5295
Db 19667 GACACCCCAACAAGCTCCAGGCCACACGCCCAAGAGAGTCTCAGGAAAGCAGACTGAA 19726
QY gaagaaatttttagcatttagaaaaaaccgcctatcagcaggcaagcccatcacacacccc 5355
Db 19727 GAAGAAATTTTAGCATTTTAGGAAACAAGGCGCATCAGCAGGCAAGGCCATGCACACACC 19786
QY aaaccacagtagtggaagagaaagacatcaaacgcgttttttgggaaactccaagtgcagaaa 5415
Db 19787 AAACCCAGCAGTAGGTGAAGAGAAAGACATCAACACGTTTTTGGGAACCTCCAGTGCAGAA 19846
QY ctgaccagcgacgagaaatttacctggcagcaatacagcgtctacaaaactcgtlaagaaaaag 5475
Db 19847 CTGGACCCAGCCAGGAATTTACCTTGGCAGCANATAGACGGCTACAAACTCTGTAAGGAANA 19906
QY gccagcgtctagaagaactgactggcttcagagagcttttccagacaccatgcactgat 5535
Db 19907 GCCCAGGCTCTAGAAGAACTGACTGGCTTCAGAGAGCTTTTCCAGACACCATGCACTCAT 19966
QY aacccacagctgatgagaaataactacaaaaatactctgcaaatctccgcgaactcagac 5595
Db 19967 AACCCACAGCTGATGAGAAAACTTACCAAAAAATACTCTGCAAACTCTCCGCAATCAGAC 20026
QY ccagcgacaccccccaacacacacaaacgagcccaagagaagcctcaagaagaagcagac 5655
Db 20027 CCAGCGACACCCCAACAACACANAGACACGGCCCAAGAGAGCCCTCAGAAAGCAGAC 20086
QY gtagagaaagaatttttagcattcaggaactaacacacatcagcaggcaagccatgcac 5715
Db 20087 GTAGAGGAAGAAATTTTAGCATTTTCCAGGAAACTAACACCATCAGCAGGCAAGGCCATGCAC 20146
QY acgcctaaagcagcagtagtggaagagaaagacatcaacacatttctggggactccagtg 5775
Db 20147 ACGCTTAAGCAGCAGTAGGTGAAGAGAAAGACATCAACACATTTTGGGGACTCCAGTG 20206
QY gagaacctggacctgctaggaaatttacctggcagcagagagcgcacaaaactccctaaa 5835
Db 20207 GAGAACTGGACCTGCTTAGGAAATTTTACCTGGCAGCAAGACAGCGCCACAACTCCTTAA 20266
QY gaaaagcccaagcgtctagaagatctctggcttcaaaagctcttccagacacccaggt 5895
Db 20267 GAAAGGCCAAGGCTCTTAGAAGATCTGGCTGGCTTCAAGAGCTCTTCCAGACACACAGGT 20326
QY cacactgagaaatcaatgaccgatgacaaatcacagaagtatcctcgaatctcccaaa 5955
Db 20327 CACACTGAGGAATCAATGACCGATGACAAATCACAGAAGTATCTCTGCAAACTCTCCACAA 20386
QY ccagacccagctcaaaaccccccaacagctccaagcaacgactcaagatctccttggggaaa 6015
Db 20387 CCAGACCCAGTCAAAACCCCAACAAAGCTCCAAGCAACGACTCAAGATATCTTTGGGANA 20446
QY gtagggtgtaagaagaaggtccctaccagtcggcaagctcacacagacgctcagggaaagacc 6075
Db 20447 GTAGGTGTAAGAAGAGAGGTCTTACCAGTCGGCAAGCTCACACAGAGCTCAGGGAAGACC 20506
QY acacagacacacagagagacagcagagatggaaaagacatcaaacgctttaagaatct 6135
Db 20507 ACACAGACACACAGAGACAGCAGGAGATGGAAAGAGCATCAAGCGTTTTAAGGAATCT 20566
QY gcaagcagatgctggacccagcaaaactatggaaactggggtggagagtgggccagaaca 6195
Db 20567 GCAAAGCAGATGCTGGACCCAGCAAACTATGGAACCTGGATGGAGAGTGGCCCAAGAACA 20626
QY cctaaggaagagcccaactcactagaagacctggccggcttcaaaagagctcttccagaca 6255
Db 20627 CCTAAGGAAGAGGCCCAATCAGTAGAGAGCTGGCCGGCTTCAAGAGCTCTTCCAGACA 20686
QY ccagacacacactagagaatcaacaactgatgacaaaactaccaaaaatagcctcgaatct 6315
Db 20687 CCAGACCCACTGAGGANTCAACAACTGATGACAAAATACTACCAAAATAGCCTGCAAAATCT 20746

QY	6316	ccaccaccagaatcaatggacactccaacaaagcacaaggaggcgcccccaaacactttg	6375
Db	20747	CCACCACCAAGATCAATGGACACTCCACAAGCACAAGGAGGGGCCCAAAACACCTTTG	20806
QY	6376	gggaaaaggatatagtggaagagctctcagccctgaagcagctcacagaccacacac	6435
Db	20807	GGGAAAAGGGATATAGTGAAGAGCTCTCAGGCCCTGAAGCAGCTCACAGACCACACAC	20866
QY	6436	acagacaaagtaccaggagatgagataaaggcatacaacgctgctcagggaactgcaaaa	6495
Db	20867	ACAGACAAAGTACCAGGACATGAGGATAAAGGCATCAACGTGTTCAGGGAAGTGCAAA	20926
QY	6496	cagaactggacccagcagaagtgttaactggttagcaagaggcagccaagaactcctaag	6555
Db	20927	CAGAAACTGGACCCAGCAGCAAGTGTAACTGGTAGCAAGAGGCGACCAAGAACTCCTAAG	20986
QY	6556	ggaaaagcccaacctagaaagacttggctggtctgaaagagctcttcagacacacagta	6615
Db	20987	GGAAAAGCCCAACCCCTAGAAGACTTGGCTGGCTTGAAGAGCTCTTCAGACACCCAGTA	21046
QY	6616	tgcactgacaagccacagactcacgagaaaactaccaaataagcctgcagatctccacaa	6675
Db	21047	TGCACTGACAAAGCCACGACTCAGGAAAACTACCAAAATAGCCTGCGAGATCTCCACAA	21106
QY	6676	ccagaccagtggtgtaccccaacaattctcaaggccacagctccaagagaagctctcaggaaa	6735
Db	21107	CCAGACCCAGTGGGTACCCCAACAATCTTCAAGGCCACAGTCCAAAGAGAAGTCTCAGGAAA	21166
QY	6736	gcagacgttagaaggaagaatccttagcactcagggaagaaacgaacccatcagtaggaagct	6795
Db	21167	GCAGACGTAGAGGAAGAATTCCTTAGCACCTCAGGAAACGAAACACCACTCAGTAGGGAAAGCT	21226
QY	6796	atggacacaccacaaacccagcagagtgatgagaagacatgaagcatttatgggaact	6855
Db	21227	ATGGACACACCCAAACCCACAGAGGTGATGAGAAGACATGAAAGCATTTATGGGAAT	21286
QY	6856	ccagtcagaaaattggacctgcaggaattttacctggcagcaaaaagatggccacaact	6915
Db	21287	CCAGTGCAGAAATTTGGACCTGCCAGGAAATTTACCTGGCAGCAAAAAGATGGCCACAACT	21346
QY	6916	cctaaggaaaagcccgagctctagaagacctggctggtctcaagagagctcttcagaca	6975
Db	21347	CCTAAGGAAAAGGCCCGAGCTCTAGAAGACCTGGCTGGCTTCAAGAGACTTTCACAGCA	21406
QY	6976	ccaggcactgaagcccacgactgatgagaaaactaccaaatagcctgcgaatctcca	7035
Db	21407	CCAGGCACTGACAAAGCCACGACTGATGAGAAAACTACCAAAATAGCCTGCAAAATCTCCA	21466
QY	7036	caaccagaccagtggaaccccccagcaagcacaagcaacggcccccaagagaacctcagg	7095
Db	21467	CAACCAGACCCAGTGGACACCCCGACCAAGCACAAGCAACAGGCCCAAGAGAAACCTCAGG	21526
QY	7096	aaagcagcgttagaggaagaatttttagcactcaggaagaaacgaacacatcagcaggcaaa	7155
Db	21527	AAAGCAGACGTAGAGGAAGAATTTTATAGCACTCAGGAAACGAACACCATCAGCAGGCAA	21586
QY	7156	gcattgacaccccccaaccagcagtaagtgatgagaaaaatatcaacatttgtgga	7215
Db	21587	GCCATGGACACCCCAAAACGACGTAAGTGTATGAGAAAAATATCAACACATTTGTGGAA	21646
QY	7216	actccagtgacaaaactggacctgcttaggaattttacctggcagcagcaagagacagccacag	7275
Db	21647	ACTCCAGTGCAGAAACTGGACCTGCTAGGAAATTTTACCTGGCAGCAAGAGACAGCCACAG	21706
QY	7276	actcctaagaaaaagcctgaggtcttagaggacctggttggtctcaaaagaactcttcag	7335
Db	21707	ACTCCTTAAGGAAAAGGCTGAGGCTCTAGAGGACCTCGTGGTCTCAAGAAACTCTTCCAG	21766
QY	7336	acaccaggtcactagagaatacaatgactgatgacaaaatcacagaagtatcctgtaaa	7395
Db	21767	ACCCAGGTCACTAGGAATCAATGACTGATGACAAAATCACGAAGTATCCTGTAA	21826

QY	7396	tctccacgccagagtgatctcaaaaactcaagaagctccaagcacaaggtccaagatacc	7455
Db	21827	TCTCCACAGCCAGAGTCAATTCAAAACTCAAGAGCTCCAAAGCAAGGCTCAAGATACC	21886
QY	7456	ctggtaaaagtggacatgaagaagagagccctagcagctcaagctcacacagacatca	7515
Db	21887	CTGGTAAAGTGGACATGAAGAAGAGGCCCTTAGCAGTAGCAAGCTTCACACGACATCA	21946
QY	7516	ggggagactlacgcacacacacagagcccaacaggagatagtagagcatcaaaagcgtt	7575
Db	21947	GGGGACACTACGCAAAACACACACAGAGCCCAACAGGAGATAGTAGAGCATCAAAAGCGTTT	22006
QY	7576	agggagtctccaaagcagatctctggccccagcagcaagtgtaactggttagcagagagcag	7635
Db	22007	AAGGAGTCTCCAAAGCAGATCTCTGGACCCACAGCAGCAAGTGAATCTGGTAGCAGGAGCAG	22066
QY	7636	ctgagaactcgttaagaaaaagcccgctgctctagaagacctggttgacttcaagagctc	7695
Db	22067	CTGAGAACTCGTAAAGAAAAGGCCCGTGTCTTAGAAGACCTGGTTGACTTCAAGAGACTC	22126
QY	7696	tctcagcaccaggtcacactgaaggtcaatgactattgacaaaaacacaaaaattccc	7755
Db	22127	TTCTCAGCACCAGGTCACACTGAAGACTCAATGACTATTTCACAAAAACACAAAAATTC	22186
QY	7756	tgcaaatctccccccacagaaactaaacagacactgccacgagcacaagagatgccccaag	7815
Db	22187	TGCAAAATCTCCCCACACAGAACTTAACAGACACTGCCACGAGCACAAAGAGATGCCCAAG	22246
QY	7816	acaagtccaggaagaaagtaaaaggagagctctcagcagcttagagaggtcagcacaaca	7875
Db	22247	ACACGTCCACAGGAAAAGTAAAGAGGAGGCTCTCAGCAGTTGAGAGGCTCACGCCAAACA	22306
QY	7876	tcagggcaagcagcacacacacaaagaaccagcagcggtgatgaggcatcaaaagta	7935
Db	22307	TCAGGCGCAAGCACACACACACAAAGAACACGACGACGCGGTGATGAGGCGCATCAAGTA	22366
QY	7936	ttgaagcaacgtgcaagagaacccaaacccagtagaagagaaacccagcagagagag	7995
Db	22367	TTGAAGCAACGTCAAAAGAAAGAACCAACCCAGTAGAAGAGAAACCCAGCAGGAGAAGG	22426
QY	7996	ccaagagcacctaaagaaaagcccaacccctggaagacctggccggttcacagagctc	8055
Db	22427	CCAAGAGCACCTAAGGAAAAGGCCCAACCCCTGGAGACCTTGGCCGCTTCACAGAGCTC	22486
QY	8056	tctgaacacatcaggtcacactcaggaatcactgctcgtgcgaagccacttaaaatccc	8115
Db	22487	TCTGAACATCAGGTCACACTCAGGAATCACTGACTGCTGGCAAGCCACTAAATACCC	22546
QY	8116	tgcgaaatctccccacttagaagtggtagacaccacagcaagcaaaagagatctcag	8175
Db	22547	TGCGAAATCTCCCCACTAGAAAGTGGTAGACACCACACAGCAACAAAGAGGCATCTCAGG	22606
QY	8176	acaagtgtcagaggtacaagttaaaagaagagccttcagcagctcaagttcacacaaca	8235
Db	22607	ACACGTGTGCAGAAAGTACAAAGTAAAGAAAGAGGAGGCTTCACAGCTCAAGTTTCACAACA	22666
QY	8236	tcaggggaacaccagatgcagacaaagaaccagcaggtgaagataaaggcatcaaaagca	8295
Db	22667	TCAGGGGAAACACCGGATGCAGACAAAGAACACGACGCTGAAGATAAAGGCATCAAGCA	22726
QY	8296	ttgaaggaattctcaaaacagacacgggctccagcagcaagtgtaactggcagcagagaga	8355
Db	22727	TTGAAGGAATCTCAAAAACAGACACCGGCTCCAGCAGCAAGTGTAACTGGCAGCAGGAGA	22786
QY	8356	cgcccaagacacccagggaaaagtgcgaagccataagaacacctagctggcttcaagac	8415
Db	22787	CGCCCAAGACACCCAGGAAAAGTGCCTCAAGGCCATAGAAGACCTAGCTGCTTCAAAAGAC	22846
QY	8416	ccagcagcaggtcacactgaagaatcaatgactgatgacaaaaaacctaaaaatccctgc	8475
Db	22847	CCAGCAGCAGGTCACACTGAAGAAATCAATGACTGATGACAAAACCACTAAATACCTGCG	22906
QY	8476	aatcatcaccagaactagaagacaccgcacaagctcaaaagagacgcccagggacacgt	8535

Db 22907 AAATCATCACAGAACTAGAGACACCGCAACAAAGCTCAAGAGACGCGCCAGACACGTT 22966
Qy 8536 gccagaaaagtgaagtgaagagagctgttagcagttagcagctgcacacaaacctca 8595
Db 22967 GCCAGAAAAGTAGAAGTGAAGGAGAGAGCTGTACAGTGTGCAAGCTTCACAAACCTCA 23026
Qy 8596 ggggagaccacgcacacgacaaagagccggtagtgtagtgagggcgaagggcagaaagcattt 8655
Db 23027 GGGGAGACCACGCACACCGACAAAGAGCGGTAGGTGAGGCAAGGCACGAAAGCATTT 23086
Qy 8656 aagcaacctgcaagcggaaacgtggagcagaagaatgtaattgagcagcagagagaccca 8715
Db 23087 AAGCAACCTGCAAAGCGGAACGTGGAGCGCAGCAAGATGTAATTGGCAGCAGGAGACGCCA 23146
Qy 8716 aagcaccctaaagaaaagggcccaaccctggaagaccctggcagccttccaaagagctctct 8775
Db 23147 AGAGCCTTAAGGAAGAGGCGCAACCCCTGGAAGACCTGGCCAGCTTCCAAAGACTCTCT 23206
Qy 8776 caaacaccaggccacactagggaactggcaaatgggtgctgtgatagctttacaagcgt 8835
Db 23207 CAACACACGAGCCACACTGAGGAACCTGGCAAAATGGTGTGATAGCTTTACAAGCGCT 23266
Qy 8836 ccaagcaaacacactgacagtggaaaacctctataaaatatccagaagagttcttcgggccc 8895
Db 23267 CCAAGCAACACACCTGACAGTGGAAAAACCTCTAAAAATATCCAGAGAGTCTCTCGGGGCC 23326
Qy 8896 cctaaagttagaacccgtggagacgtgtgaagcaccagagaccctgtaaaaatcacaaagc 8955
Db 23327 CCTAAAGTAGAACCCGTGGAGAGCTGGTAAGCACACAGACACCTGTAAATATCAAAAGC 23386
Qy 8956 aagcaaacacttccttgcctccactgcccctcaagaggggaggtggcaagatggaagc 9015
Db 23387 AAAAGCAACACTTCCCTGCCCCACTGCCCTTCAAGAGGGGAGGTGGCAAGATGGAAGC 23446
Qy 9016 gtcaagggaaccaagagctgcgtgcatgccagcaccagcaggaagaattgtggaggagctg 9075
Db 23447 GTACGGGGAACCAAGAGGCTGCGCTGTCATGCCAGCACAGAGGAATTTGTGGAGGAGCTG 23506
Qy 9076 ccagccagcagaagcagagaggtgtgtcccgagggcagagggcaaatcatccgaaccctg 9135
Db 23507 CCAGCCAGCAAGAGCAGAGAGGTGTCTCCAGGCGCAGAGCAAAATCATCCGAACCCGTG 23566
Qy 9136 gtcatcatgaagagaagtttgaggaactcttgcataaagaattgaacctgcgggaagctg 9195
Db 23567 GTCATCATGAAGAGAGAGTTTGTGAGACTTCTGCAAAAAGAAATTGAACCTGCGGAAGAGCTG 23626
Qy 9196 acagcaacacatgaacaacaaacaaagagagacacaaatataagactcgttcctctgaa 9255
Db 23627 AACAGCAACGACATGAACCAACCAAGAGGAGACACAAATTTACAAGACTCGGTCCCTGTAA 23686
Qy 9256 aataagg 9262
Db 23687 AATAAGG 23693

RESULT 6
AL390236/c
LOCUS Human DNA sequence from clone RP11-380J17 on chromosome 10,
DEFINITION complete sequence.
ACCESSION AL390236
VERSION AL390236.21 GI:17381338
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 123057)
Sycamore, N.
Direct Submission
Submitted (30-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

COMMENT humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Dec 5, 2001 this sequence version replaced gi:17017713.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TrEMBL; Wp., WormPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-380J17 is from the library RPCR-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3 6
IMPORTANT: This sequence is not the entire insert of clone
RP11-380J17 It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP11-380J17 is at 123057 in this
sequence. The true left end of clone RP11-85C15 is at 103775 in
this sequence. The true right end of clone RP11-4C20 is at 2000 in
this sequence.
FEATURES
source
1. 123057
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-380J17"
/clone_lib="RPC1-11.2"
BASE COUNT 30696 a 29821 c 29797 g 32743 t
ORIGIN
Query Match 53.5%; Score 5227; DB 9; Length 123057;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 5777; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 3475 gaattcttagcactcaggaactaacaccatcagcagggaaagccatgcttaacgcccacaa 3534
Db 123057 GAATTCCTAGCCTCAGGAAACTTAACACCATCAGCAGGAAAGCCATGCTTACGCCCAA 122998
Qy 3535 ccagcaggaggtgatgagaagacattaaagcatttattggaaacctccagtcgagaaactg 3594
Db 122997 CCAGCAGGAGGTGATCAGAAAGACATTTAAAGCATTTTGGGAACCTCCAGTCGAGAAACTG 122938
Qy 3595 gacctggcagaactttacctggcagcaaaagacagctacagactcctcaggaagagccc 3654
Db 122937 GACCTGGCAGGAACCTTTACCTGGCAGCAAAAGACAGCTACAGACTCTCTTAGGAAAGGCC 122878
Qy 3655 caggctctagaagacacctggctggcttttaaagagctctctccagactcctggccacacgag 3714
Db 122877 CAGGCTCTAGAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGAGCTCTTCCAGACCTGTCACACCGAG 122818
Qy 3715 gaattagtgtcgtgtgtaaaacacactaaaaataccctgcgactctccacagtcagaccca 3774
Db 122817 GAATTAGTGTGCTGTGTTAAAGCACTTAAATAACCTTCGCGACTCTCCACAGCTCAGACCCA 122758
Qy 3775 gtggaacccccacaagcacaagaacacccacacacacagagagatgattcaggaagacagatga 3834
Db 122757 GTGACACACCCCAACAGCAAAAGCAACGACCCCAAGAGAGTATCAGGAAGCAGATGTA 122698

Qy 3835 gagggagaactcttagctgaggaatactaatgcatcagcaggcaaaagccatgcacacg 3894
Db 122697 GAGGGAGAATCTTATAGCTGCAGGAATCTAATGCTCATCAGCAGGCAAGCCATGCACAGC 122638
Qy 3895 cctaaaccatcagtaggtgaagagaaagacatcatatttttggaactccagtcag 3954
Db 122637 CCTAAACCATCAGTAGGTGAAGAGAAAGACATCATATTATTTGGGAACTCCAGTGAG 122578
Qy 3955 aaactggacctgacagagaacttaaccggaagcagcaagagagcgcacaaactcctaagaa 4014
Db 122577 AAACCTGGACCTGACAGAACTTAACCGGCAAGAGACGGCCCAAACTTCCTAAGGAA 122518
Qy 4015 gagggccaggtcttgaagacctgactggctttaagagctcttccagacccttgatcat 4074
Db 122517 GAGGCCAGGCTCTGGAAGACCTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTCAT 122458
Qy 4075 actgaagaagcagtggtgctggtggcaaaactactaaatgacctggaattcttccaca 4134
Db 122457 ACTGAAGAAGCAGTGGCTGCTGGCAAACTACTAAAATGCCCTGGGAATCTTCTCCACCA 122398
Qy 4135 gaatcagcagacaccccccaagcacagaagcagcccaagacacacctttggaagaaagg 4194
Db 122397 GAATCAGCAGACACCCCCCAAGCAAGAGGCGCCCAAGACACCTTTTGGAGAAAGG 122338
Qy 4195 gacgtacagaagaggtctctcagccctgaagagctcacacagacatcaggggaaaccaca 4254
Db 122337 GACGTACAGAAGGAGCTCTCAGCCCTGAAGAAGCTCACACAGACATCAGGGGAAACCCACA 122278
Qy 4255 cacacagataaagtaccagggtgaggaataaaagcataacgcgttttagggaaactgca 4314
Db 122277 CACACAGATAAAGTACCAGGAGGTGAGGATAAAGCATCAACGCGCTTTTAGGGAACTGCA 122218
Qy 4315 aaacagaactggacccagcagcagtgtaactggtgaagagcagcccaaaactaag 4374
Db 122217 AAACAGAACTGGACCCAGCAGCAAGTGTAACTGGTACAGAGGACCCCAAACTTAG 122158
Qy 4375 gaaaaggcccaaccccttagaagacctggtggtggaagagctcttccagacaccagta 4434
Db 122157 GAAAAGGCCAACCCCTTAGAAGACCTGGCTGGCTTGAAGAGCTCTTCCAGACACCACTA 122098
Qy 4435 tgactgcaagcccaagactcacgagaaactaccatacgaatagcctgcagatcacaca 4494
Db 122097 TGCACTGCAAGCCCAACGACTCACGAGAAACTACCAAAATAGCCTGCAGATCACAACTCA 122038
Qy 4495 gaccagtggaacacacccaagctcccaagccacagtcacagagagctcaggaagtg 4554
Db 122037 GACCCAGTGGACACACCAACAGCTCCAAAGCCACAGTCCAAAGAGAGTCTCAGGAAAGTG 121978
Qy 4555 gacgtagaagaagaattcttcgactcaggaacacgcaaccatcagcaggcaaaagccatg 4614
Db 121977 GACGTAGAAGAAGAATTCTTGCACTCAGGAACCAACACCATCAGCAGGCAAAAGCCATG 121918
Qy 4615 cacacacccaaacacgagtagtggtgagaaaaacatacagactttatgggaactcca 4674
Db 121917 CACACACCCAAACCAACGACTAGTGGTGGAGAAAAACATCTACGCAATTTATGGGAATCTCA 121858
Qy 4675 gtgcagaactggacctgacagagaacttaactggcagagagcagcgtacaaactcct 4734
Db 121857 GTGCAGAACTGGACCTGCACAGAGAACTTAACCTGGCAAGAGAGCGCTACAAACTCCT 121798
Qy 4735 aaggaaaaggccaggtcttagaagacctggtggtttaaagagctcttccagacacga 4794
Db 121797 AAGGAAAAGGCCAGGCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACACGA 121738
Qy 4795 ggtcactgaggaatactaactaacgataaaaaactgcaaaagtgcctgcaaatctcca 4854
Db 121737 GGTCACTAGGGAATCAATGACTAACGATAAAACTGCGAAAGTAGTACCTGCAAACTCTTCA 121678
Qy 4855 caaccagacctagacaaaacccagcagctccaagcagcagcgtcaaacatccctggg 4914
Db 121677 CAACCAGACCCAGACAAAACCCAGCAAGCTCCAAAGCAGCGGCTCAAGACATCCCTGGGG 121618

Qy 4915 aaagtggcggtgaagaagagctcctagcagttggcaagctcacacagatcagagag 4974
Db 121617 AAAGTGGCGTGAAGAAGAGAGCTCTAGCAGTTGCAAGCTCACACAGACATCAGGAGAG 121558
Qy 4975 actacacacacacacagagcccaagagagatggttaagagcatgaaagcatttatggag 5034
Db 121557 ACTACACACACACACAGAGGCCAACAGAGATGTTAAGAGCATGAAAGCATTTATGGAG 121498
Qy 5035 tctccaaagcagatcttagactcagcagcaagtctaaactggcagcaagagcagctgaga 5094
Db 121497 TCTCCAAAGCAGATCTTAGACTCAGCAGCAAGCTTAAGTGGCAGCAAGAGGAGCTGAGA 121438
Qy 5095 actcctaagggaagtctgaagtccctgaagacctggcggcttcatcgagctcttccag 5154
Db 121437 ACTCTAAAGGAAAGTCTGAAGTCCCTGAAGACCTGGCGGGCTTCATCAGCTCTCCAG 121378
Qy 5155 acaccaagtccacactaaggaataactgactaatgaaaaaactaccaaaagtatcctacaga 5214
Db 121377 ACACCAAGTCACTAAGGAATCAATGACTAACGAAAAAACTACCAAAAGTATCCTACAGA 121318
Qy 5215 gtttcacagcagaccttagtgacaccccccaacaaagctccaagccacagcccaagagaagt 5274
Db 121317 GCTTCACAGCCAGACCTAGTGGACACCCCAACAGCTCCAGCCACAGCCCAAGAGAGCT 121258
Qy 5275 ctcaggaagcagacactgaagaagaatttttagcattttaggaacaaacgcacatcagca 5334
Db 121257 CTCAGGAAAGCAGACACTGAAGAAGAAATTTTAGCATTTTAGGAAACAAACGCCATCAGCA 121198
Qy 5335 ggcaaaagccaatgcacacacccccaaacagcagtaggtgaagagaaagacatacaacgcttt 5394
Db 121197 GGCAAAAGCCATGCACACACCCCAACAGCAGTAGGTGAAGAGAAAGACATCAACAGCTTT 121138
Qy 5395 ttgggaactccagtcgagaaactggaccagccaggaataattccttggaagcaataagacgg 5454
Db 121137 TTGGGAACCTCAGTCGAGAAACTGGACCCAGCCAGGAAATTTACCTGGCAGCAATAGACCG 121078
Qy 5455 ctacaactcgttaaaggaaagcccaggctctagaagaactgactggtctcagagagctt 5514
Db 121077 CTACAAACTCGTAAGGAAAAAGGCCAGGCTCTAGAAGAACTGACTGGCTTCAGAGAGCTT 121018
Qy 5515 ttccagacacatgactgataacccccacagctgatgagaaaaactacaaaaatacttc 5574
Db 121017 TTCCAGACACCATGACTGATTAACCCACAGCTGATGAGAAAACTACCAAAAAATACTCTC 120958
Qy 5575 tqcaaatctccgcaatcagacccagcggacaccccaacaaacacaaagcaacgagcccaag 5634
Db 120957 TGCAAACTCTCCCATCAGACCCAGCGGACACCCCAACAAACACAAAGCAACGGCCCAAG 120898
Qy 5635 agaagcctcaagaagcagacgttagaggagaagaatttttagcattcagggaactaaccacca 5694
Db 120897 AGAAGCCTCAAGAAAGCAGACAGCTAGAGGAAGAAATTTTAGCATTCAGGAAACTAACACCA 120838
Qy 5695 tcagcaggcaaacatgcacacgccttaagcagcagtaggtgaagagaaaagacatcaac 5754
Db 120837 TCAGCAGCCAAAGCCATGCACACGCCTAAAGCAGCAGTAGTGGTGAAGAGAAACATCAAC 120778
Qy 5755 acatttgtgggactccagtgagaaactggacctgctaggaatttacctggcagcaag 5814
Db 120777 ACATTTGTGGGACTCCAGCTGGAGAAACTTGACCTGCTAGGAAATTTACTTGGCAGCAAG 120718
Qy 5815 agacggccacaaactcctaaagaaaggccaaggtcttagaagatctggctggcttcaaa 5874
Db 120717 AGACGCCCAAAACTCCTAAAGAAAAAGGCCAAGGCTCTAGAAGATCTGGCTGGCTTCAAA 120658
Qy 5875 gagctcttcagacacccaggtcacactgaggaataataagaccgatacaaaaaatcacagaa 5934
Db 120657 GAGCTCTTCCAGACACCCAGGTCACTGAGGAATCAATACCCGATGACAAAAATCACAGAA 120598
Qy 5935 gtatcctgcaaatctccacaacacagccagctcaaaaccccccaagctcccaagcaacga 5994
Db 120597 GTATCCTGCAAACTCTCACAAACAGACCCAGCTCAAAAACCCCAACAGCTTCCAAAGCAACGA 120538
Qy 5995 ctcaagatatacttggggaaagtaggtgtaagaagaggttctaccagtcggaagctc 6054

|||||CTCAAGATATCCCTTGGGGAAGTAGTGTTGAAAGAGAGAGTCTTACCAGTCGGCAAGCTC 120478
QY 6055 acacagacgtcaggggaagacacacagacacacagagagacagcagagatggaagagc 6114
Db 120477 ACACAGACGTCAGGGAAGACACACAGACACACAGAGAGACACAGGAGATGGAAGAGC 120418
QY 6115 atcaagcgctttaaggaatctgcaagcagatgctggaccagcaaaactatggaactggg 6174
Db 120417 ATCAAGCGCTTTAAGGAATCTGCAAGCAGATGCTGGACCAGCAAACTATGGAACCTGGG 120358
QY 6175 atggagagtgccagaagaaacacctaaggaagagggcccaatcaatagaaagacctggccggc 6234
Db 120357 ATGGAGAGGTGGCCAGAACACCTTAGGAAGAGGGCCCAATCACTAGAAAGACCTTGGCCGGC 120298
QY 6235 ttcaagagagctctccagacacacagacacacactgaggaatcaacaactgatgacaaaact 6294
Db 120297 TTCAAGAGAGCTCTCCAGACACACAGACACACACTGAGGAATCAACAACCTGATGACAAAAC 120238
QY 6295 accaaaatagctgcgaatctccaccaccagaatcaatggacactccaacaagcacaagg 6354
Db 120237 ACCAAAATAGCTTGCAAATCTCCACCACCAGAATCAATGGACACTCCAACAAGCACAAGG 120178
QY 6355 agcgcccaaaaacacctttggggaagggatatagtggaagagctctcagccctgaag 6414
Db 120177 AGCGGGCCCCAAAACACCTTTGGGGAAGGGATATAGTGGAGAGGCTCTCAGCCCTGAAG 120118
QY 6415 cagctcacagacacacacacagacaaagtaccagagagatgagataaaggatcaac 6474
Db 120117 CAGCTCACAGACACACACACAGACAAAGTACCAGGAGATGAGGATTAAGGCATCAAC 120058
QY 6475 gtgttcagggaactgcaaaacagaactggaccgcagcagcaagttaactggtagcaag 6534
Db 120057 GTGTTCAGGGAACCTGCAAAACAGAACTGGACCCAGCAAGTGTAACTGGTGTGACAAG 119998
QY 6535 agcgaccaagaactcctaagggaagaaagcccaaccctcagaagacttggctgtgaaa 6594
Db 119997 AGCGACCAAGAAGTCTCTAAGGGAAGGCCAACCCCTTAGAAGACTTGGCTGGCTTTGAA 119938
QY 6595 gagctctccagacacagtatgactgacaaagccacagactcacgagaaaactaccaaa 6654
Db 119937 GAGCTCTCCAGACACCAATATGACTGTACAGACCCACGACGCTCATGAGAAAATACCAAA 119878
QY 6655 atagctgcgaatctccacaacacagaccagctgggtacccccacaatcttcaagccacag 6714
Db 119877 ATAGCTGCAGATCTCCACAACCAGACCCAGTGGGTACCCCAACAATCTTCAAGCCACAG 119818
QY 6715 tccaagagaagctcaggaagcagacgtagaggaagaatccttagcactcaggaaacga 6774
Db 119817 TCCAGAGAAGTCTCAGGAAGCAGACGTTAGGGAAGAATCCTTAGCCTCAGGAAACGA 119758
QY 6775 acaccatcagtaggggaagctaTggacacaccccaaccagcagaggtgtgatgagaagac 6834
Db 119757 ACACCATCAGTAGGGAAGCTATGGACACACCCCAACACAGCAGGAGGTGATGAGAAGAC 119698
QY 6835 atgaagcatttatgggaactccagtcgagaattggacctgcaggaatttacctggc 6894
Db 119697 ATGAAAAGCATTTATGGGAACCTCCAGTGCAGAAATTTGGACCTGCCAGGAAATTTACCTGGC 119638
QY 6895 agcaaaagatggccacaactcctaaggaaaagggccaggtccttagaaagacctggctggc 6954
Db 119637 AGCAAAAGATGGCCCAAACTCTCTAAGGAAGAGGCCACCGCTCTAGAAGACCTTGGCTGGC 119578
QY 6955 ttcaagagactcttcagacacacagcactgacaagcccaaccagactgatgagaaaactacc 7014
Db 119577 TTCAAGAGAGCTCTCCAGACACACAGCAGCTGACAGACCCACGACTGATGAGAAAATACC 119518
QY 7015 aaaaatagctgcgaatctccacaacacagaccagctggacacccccagcaagcacaagcaa 7074
Db 119517 AAAATAGCCTGCAAAATCTCCACAACACAGACCCAGTGGACACCCCAAGCAAAAGCAA 119458
QY 7075 cggcccaagaaacctcaggaaagcagacgttagaaggaagaatttttagcactcaggaaa 7134
|||||

Db 119457 CGCCCCAAGAGAAACCTCAGAAAGCAGAGCTAGAGGAAGAAATTTTATGACACTCAGGAAA 119398
QY 7135 cgaacaccatcagcagcagaaagccatggacaccccccaaacccagcagtaagtgtatgagaaa 7194
Db 119397 CGAAACACCATCAGCAGGCAAGCCATGGACACACCAACCAACCAAGCTAAGTGATGAGAAA 119338
QY 7195 aatatcaacacatttgtggaactccagtcgagaactcagcctgcttaggaaatttacct 7254
Db 119337 AATATCAACACATTTGTGGAACCTCCAGTGCAGAAACTGGACCTGTAGCAAAATTTACCT 119278
QY 7255 ggacgaagaagacagccacagactcctaaggaagaaagctgaggtccttagaggacctgggt 7314
Db 119277 GGCAGCAAGAGACAGACACAGACTCTTAAGGAAAAGGCTGAGGCTCTAGAGGACCTGGTT 119218
QY 7315 ggtttcaagaactctctccagacacacaggttcacactgaggaatcaatgactgatgacaaa 7374
Db 119217 GGGTTCAAGAACTCTTCCAGACACCCAGGTACACTGAGGAATCAATGACTGATGACAAA 119158
QY 7375 atcacagaagtaTcctgttaaatctccacagccagagtcattcaaaaacctcaagaagctcc 7434
Db 119157 ATCACAGAAGTATCTGTAAATCTCCACAGCCAGAGTCAATTCAAAACCTCAAGAAGCTCC 119098
QY 7435 aagcaagcgctcaagataccccctgtggaagtggacatgaaagaagggcccccttagcagtc 7494
Db 119097 AAGCAAAGGCTCAAGATACCCCTGCTGTAAGTGGACATGAAGAAGAGGCCCTTAGCAGTC 119038
QY 7495 agcaagctcacacgacatcaggggagactacgcaaacacacacagagcccaacagagat 7554
Db 119037 AGCAAGCTCACAGGACATCAGGGGAGACTACGCCAAACACACACAGAGCCACAGAGAT 118978
QY 7555 agtaagagcatcaaaagcgtttaaggagctccaaaagcagatccttgaccaccagcagcaagt 7614
Db 118977 AGTAACAGCATCAAGCGTTTAAAGGAGTCTCCAAGACAGATCTCGACCCAGCAGCAAGT 118918
QY 7615 gtaactggtagcagaggcagctgagaactcgttaaggaaaaagggccgctccttagaagac 7674
Db 118917 GTAACCTGGTAGCAGGAGGAGCTGAGAAGTCTGTAAGGAAAAGGCCGCTGCTCTAGAAGAC 118858
QY 7675 ctggttgacttcaaaagctctcttcagcagcaggttcacactgaagagtcataagctatt 7734
Db 118857 CTGGTTGACTTCAAAAGAGCTCTCTCAGCACCCAGGTTCACACTGAAGAGTCAATGACTATT 118798
QY 7735 gacaaaaacacaaaaattccctgcgaattctccccaccagaaactaacagacactgcacag 7794
Db 118797 GACAAAACACAAAAATTCCTGCAAAATCTCCCCACCAGAACTAAACAGACACTGCCACG 118738
QY 7795 agcacaaagagatgccccaaagacagctccaggaagaagtaaaagaggagctcctcagca 7854
Db 118737 AGCACAAAGAGATGCCCCCAAGACACCTGCCAGGAAAGAGTAAAGAGGAGCTCTCAGCA 118678
QY 7855 gttgagagctcacgcaaacatacaggggcaagcacacacacacacacacacacagcaagc 7914
Db 118677 GTTGAGAGGCTCAGCAAAACATCAGGGCAAGGACACACACACACACAAAGAACCAAGC 118618
QY 7915 gbtgatggggcatcaaatatttgaaagcaagtcgcaaaagaaagaaacccacagtagaa 7974
Db 118617 GGTGATGAGGGCATCAAGTATTGAAAGCAACGTGCAAAAGAAACCAAAACCCAGTAGAA 118558
QY 7975 gaggaaacccagcaggaagaagccagagcacctaaggaagaaagcccaacccctggagac 8034
Db 118557 GAGGAACCCAGCAGGAGAAGGGCCAGAGCACCTTAAGGAAAAGGCCCAACCCCTTGGAAAGC 118498
QY 8035 ctggccggcttcacagagctctctgaacatcaggttcacactcaggaatcaactgactgct 8094
Db 118497 CTGGCCGGCTTCACAGAGCTCTTGAAACATCAGGTTCACACTCAGGAATCACTGACTGCT 118438
QY 8095 ggcaagccactaaaaataccctgcgaattctccccactagaagtgttagacaccacagca 8154
Db 118437 GGCAAGGCCACTAAAAATACCTTCCGAAATCTCCCCCACTAGAAGTGGTAGACACACAGCA 118378
QY 8155 agcacaaagagcactctcaggacagctgtgcagaaggtacaagttaaaagaagccttca 8214
Db 118377 AGCACAAAGAGGCATCTCAGGACACGCTGTGCAGAGGTACAAGTAAAGAAAGAGCCTTCA 118318

QY	8215	gcagtcgaagtccaacaaacatcaggggaaacccacggtgagacaaagaaacccagcgagt	8274
Db	118317	GCAGTCAAGTTCACAAAAATCATCAGGGAAACACCGATGACGACAAAGAACCCAGCAGGT	118258
QY	8275	gaagataaaggcatcaaaacattgaagaatactgcaaaacagacacgcgtccagcagca	8334
Db	118257	GAGATAAAGGATCAAAACATTGAAGAACTCTGCAAAACAGACACCGGCTCCAGCAGCA	118198
QY	8335	agtgtactggcagcaggagagcgcgaagagcaccocagggaagtgcaccaagccatgaa	8394
Db	118197	AGTGTAACTGGCAGCAGGAGACGGCCAAAGAGCACCCAGGAAAGTGCCCAAGCCATGAA	118138
QY	8395	gaactagctggcttaaaagaccacagcagcaggtcacactggaagaatacaatgactgatgac	8454
Db	118137	GACCTAGCTGGCTTCAAGACCCAGCAGCGTCACTGAAGAAATCAATGACTGATGAC	118078
QY	8455	aaacaccataaaatcccttgcaaatcatcaccagaaactagaagacaccgcacaagctca	8514
Db	118077	AAACCACATTAAATACCTTGCAAAATCATCCAGAACTTAGAAGACACCGCCAAACAGTCA	118018
QY	8515	aagagacggcccgacacgctgccagaaagtgaagtgaaggaggagctgttagcagctt	8574
Db	118017	AAGAGACGGCCAGGACAGCTGCCAGAAAGTAGAAGTGAAGGAGGAGCTCTTAGCAGTT	117958
QY	8575	ggcaactcacaaaaacctcaaggagagaccagcacaccgacaaaagacggtaggtgag	8634
Db	117957	GGCAAGCTCACAAACCTTCAGGGGAGACCGACACCGCAACAAAGACCGGTAGGTGAG	117898
QY	8635	ggcaaggccagaaagcatttaagcaacctgcaaaagcggaacgtggacgcagaagatgta	8694
Db	117897	GGCAAGGCACGAAAGCATTTAAGCAACCTGCAAGCGGAGAGCTGGAGCGCAGAGATGTA	117838
QY	8695	attggcagcagagagacagccaagagcaactaaggaaaaagggccccacccctggaagacctg	8754
Db	117837	ATTGGCAGCAGGAGACAGCAAGCAAGACACTTAAGGAAAAAGGCCAACCCCTGGAAGATCTG	117778
QY	8755	qccagcttccaagagctctctaaacaccaggccacactggaactggcaaatggtgct	8814
Db	117777	GCAGCTTCCAAGAGCTCTCTCAACACCAGCGCCACACTGAGGAACCTGGCAATGGTGCT	117718
QY	8815	gctgatagctttacaagcgctccaaagcaaacacactgacagtggaaaacctcttaaaaaa	8874
Db	117717	GCTGATAGCTTTACAAAGCGCTCCAAAGCAAAACACCTGACAGTGGAACACCTTAATAA	117658
QY	8875	tcagaagaggttcttggcgccctcaagttagaaccctggtgagacgtgtaagcaccaga	8934
Db	117657	TCAGAAGAGTTCCTTGGCGCCCTAAAGTAGAACCCGTGGAGACGCTGGTAAGCACCAGA	117598
QY	8935	gacctgtaaaaatcacaagcaaaagcaacacttccctgcccccaactgaccttcaagagg	8994
Db	117597	GACCTGTAAATCACAAGCAAAAGCAACACTTCCCTGCCCCCTCCCTTCAAGAGG	117538
QY	8995	ggaggtggcaaatggaagcgtcacgggaaccaagagcgctgcgtcatgccagcacc	9054
Db	117537	GGAGGTGGCAAAAGATGGAAGCGTCAACGGGAACCAAGAGGCTGCGCTGCATGCCAGCACA	117478
QY	9055	gaagaaattgtgaggagctgcagcccgcaagaagacagagggttgctcccgagggaaga	9114
Db	117477	GAGGAATTTGTGAGGAGCTGCCAGCCAGCAAGAGAGAGAGGGTTGCTCCCGAGGCAAGA	117418
QY	9115	ggcaaatcatcgaaacccgtggtcatctgaagaaagtttgaggacttctgcgcaaaaaga	9174
Db	117417	GGCAANTCATCCGACCCGTGCTCATCATGAGAGAAGTTTGAGGACTCTTGCAAAAAGA	117358
QY	9175	attgaacctgcggaagagctgaacagcaacgacatgataaaccacaaagagagacacaaa	9234
Db	117357	ATTGAACCTCGGAGAGCTGAACAGCAACGACATGAAACCAACAAAGAGAGAACACAAA	117298
QY	9235	ttacaagactcgttccctgaaataagg	9262
Db	117297	TTACAAGACTCGTCCCTGAAAAATAGG	117270

QY 2416 ggaggaaatgtgttcttcagtgcacagaaatgcagcgaacaaacagccatctgataaatgtctt 2475
D 3059 GGAGGAAATGTGTCTTTCAGTGCACAGAAATGCAGCAAAACAGCCATCTGATATAATGTCTT 3000
QY 2476 gcaagccctcccttaagacagcgagtgatgcagagaaaaatggaacgtagcaaaaaagccc 2535
D 2999 GCAAGCCCTCCCTTTAAGACGGCAGTGTATTAGAGAAATGGAAACGCTAGCAAAACGCC 2940
QY 2536 aggaacacctacaaaaatgactctcttgagacaaaaaacttcagatactgcagacagagcct 2595
D 2939 AGGAACACCTTACAAATGACTTCTCTGGACACAAAACCTTCAGATACTGCAGACAGCC 2880
QY 2596 tcaaaaacagtatccactgtaaacaggtcagggaaggtctacagaggttcaggaatatcacg 2655
D 2879 TCAAAAACAGTATCCACTGCAACAGCTCAGGAAGGTCTACAGAGTTCAGGAATATACAG 2820
QY 2656 aagctacctgtgaaagtgaagatgaagaaacaaaaatacagaaaattgttgagtcacacta 2715
D 2819 AAGCTACTCTGTGAAAGTAAAGTGAAGAAACAAATACAGAAATTTGTGAGTGCATCCTA 2760
QY 2716 aaaaaggttcagaaagcaactactactacaacaagagagagagagatgaaggaata 2775
D 2759 AAAAGAGTTCAGAAAGCAACACTACTACACAACAAGAGAGAGAGATGAAGGAATA 2700
QY 2776 gaaagaccttttgagacataaaggaaaaatttgaattaaaaaagaaacgataaaaaagatg 2835
D 2699 GAAAGACCTTTTGAGACATATAAGGAAATATTGAATTTAAAGAAAAACGATGAAAGATG 2640
QY 2836 aagcaaatgaagagataaagaacttggggcagaaaaatgtcaccaaatgtctgacctgaca 2895
D 2639 AAAGCAATGAAGAGATCAAGAATTTGGGGCAGAAATGTGCACCAATGTCTGACCTGACA 2580
QY 2896 gacctaaagacttgcctgtatcacagaactcatgaagacacgcacgtgcccagaaatctc 2955
D 2579 GACCTCAAGAGCTTGCTGTATACAGAACTCATGAAGACACGCGACGTGGCCAGAATCTC 2520
QY 2956 ctccaaacccaagatcatgcacaggcaccaaaagagtgagaaggcaaaaatcactaaaaatg 3015
D 2519 CTCCAAAACCAAGATCATGCAAGGCACCAAGAGTGAAGAGTGAAGGCAAAATCACTAAATG 2460
QY 3016 cctgcagctcattacaaccagaaaccaataaacacccccacacacacacacacacagttg 3075
D 2459 CCTGCCAGTCTATTACAAACAGAACCAATAAACCACCCCAACACACACAAACACAGTTG 2400
QY 3076 aaggcatccctggggaaagtagtgtgaagaagagctccttagcagtcggcaagttcaca 3135
D 2399 AAGCATCCCTGGGAAAGTAGTGTGAAGAAGAGCTCCTACAGTTCGCAAGTTTACA 2340
QY 3136 cggagctcaggggagaccacgcacagcagacagagagccagcagagatggcaagagcattc 3195
D 2339 CGGACGCTCAGGGGAGACCCAGCCACAGCCACAGAGAGCCAGCAGGAGATGGCAAGAGCATC 2280
QY 3196 aagaacttaagagtgctccaagcagataccttgagccccagcagcccggttaactggaatg 3255
D 2279 AGAACGTTTAAGGAGTCTCCAAAGCAGATCCTTGAGCCCCAGCAGCCCGTGTAACTGGAAATG 2220
QY 3256 aagaagtggccaagaacgcttaaggaaagagcgccagtcactagaagacacctggctggcttc 3315
D 2219 AAGAAGTGGCCANAGACGCCCTAAGGAAGAGGCCCTACTACTAGAAGACCTTGGCTGGCTTC 2160
QY 3316 aagagactcttcagacacaccaggtccctctgaggaatcaatgatgataaagaaactacc 3375
D 2159 AAAGAGGCTCTTCAGACACACCAGGTCCCTCTGAGGAATCAATGACTGATGAGAAAACTACC 2100
QY 3376 aatatagctgcaaaatctccaccacagaatcagtgaactccaacacagacacaaagcaa 3435
D 2099 AAATAGCTTGCRAAATCTCCACCACCAGAAATCAGTGGACACTCCAAACAGACAAAAGCAA 2040
QY 3436 tggcctaagaagaagttctcaggaagcagatgtagagaagaatctcttagcactcagggaaa 3495
D 2039 TGGCCTTAAGAGAAGTCTCAGGAAGCAGATGTAGAGGAAGAATTTCTTAGCACTCAGGAAA 1980

QY 3496 ctaacacca tcagcagggaagccatgtcttacgccccaaacagcagaggtgatgagaaa 3555
D 1979 CTAAACCATCAGCAGGGAAGGCATGCTTACGCCAAACAGCAGGAGGTGATGAGAAA 1920
QY 3556 gacattaaacatttatggaaactccagtgcagaaactggacctggcagaaactttacct 3615
D 1919 GACATTAAAGCATTTATGGGAACCTCCAGTGCAGAAACTTGACCTGGCAGGAACCTTACCT 1860
QY 3616 ggagcaaaaagacagctcacagactccttaaggaaaaagccccaggtccttagaagaccttgct 3675
D 1859 GGCAGCAAAAGACAGCTACAGACTCTCTAAGGAAAAGGCCCGCAGGCTCTAGAAAGACCTGGCT 1800
QY 3676 ggttttaagagagctcttcagactcctgtg tcacacogaggaattagtggtgtgtgttaaa 3735
D 1799 GGGTTTTAAAGAGCTCTTCCAGACTCCTGGTTCACACGAGGAATTAGTGGCTGTGGTAAA 1740
QY 3736 accaactaaaataccctgcagactctccacagtgcagaccagtgcacaccccccaacaaagcaca 3795
D 1739 ACCACTAAAATACCTTCGACTCTCCACAGTTCAGACCCAGTGGACACCCCAACAAAGCACA 1680
QY 3796 aagcaacgacccccaaagaaatatacagaaagcagatgtagagggagaactcttagcgtgc 3855
D 1679 AAGCAACGACCCCAAGAGAGTATCAGAAAGCAGATGTAGAGGGAGAACTCTTTAGCGTGC 1620
QY 3856 aggaatctaagtccatcagcaggcaagccatgcacacgcctaaacacatcagtagtgaa 3915
D 1619 AGGAATCTAATGCCATCAGCAGGCAAAAGCCATGCACACGCCCTAAACCATCAGTAGGTGAA 1560
QY 3916 gagaagacatactcatttgggaactccagtgcagaaaaactggacctgcagagaaac 3975
D 1559 GAGAAAGACATCATATATTGTGGGAACCTCCAGTGCAGAAACTTGGACCTGCAGAGAAAC 1500
QY 3976 ttaaccggcagcagaagagcgccacaaactcctaaagaagagccccaggtcctcggaaagac 4035
D 1499 TTAACCGCGCAGCAAGAGCGGCCACAACTCTTAAGGAAGAGCGCCAGGCTCTGGAAGAC 1440
QY 4036 ctgactggcttttaagagctcttcagacccctcgttcactgaaagaagcagtggtgtgt 4095
D 1439 CTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTCATACTGAAGAAGCAGTGGCTGTCT 1380
QY 4096 gcaaaactactaaaatgccctgcgaactcttcaccagaaatcagcagacaccccaaca 4155
D 1379 GGCAAAACCTACTAAAATGCCCTCGGAATCTTCCACCAGAATCAGCAGACACCCCAACA 1320
QY 4156 agcacaagaagcagccccaaagacacctttggagaaaagggacgtacagaaaggagctctca 4215
D 1319 AGCACAAAGAGCGACCCCAAGACACCTTTTGGAGAAAAGGGACGCTACAGAAGAGCTCTCA 1260
QY 4216 gcctgaagaagctcacagacatacagggggaacacacacacagataaagtaccagga 4275
D 1259 GCCTTGAAGAAGCTCACAGACATCAGGGGAAACACACACACAGATAAAGTACCAGGA 1200
QY 4276 ggtgagataaaagcatcaacgcgttttaggaactgcataacagaaactggaccacaga 4335
D 1199 GGTGAGGATAAAAGCATCAACGCGTTTAGGGAACCTGCAAAACAGAACTTGGACCCAGCA 1140
QY 4336 gcaagtgttaactgttagcaagagcgccccaaaaactcaaggaaaagccccaaaccttagaa 4395
D 1139 GCNAAGTGAACCTGTAGCAAGAGCGACCCCAAAACTTAAGGAAAAGGCCCAACCCCTTAGAA 1080
QY 4396 gacctggctggctggaagagctcttcagacaccccgatgtgcactgacaagccccagact 4455
D 1079 GACCTGGCTGGCTTGAAGAGCTCTTCCAGACACCACTAGTATGCATGACAGGCCAGGACT 1020
QY 4456 caggaataaactaccataatagctcagatcacacccagacccagtcgacacacaca 4515
D 1019 CAGGAGAAACTTACCAAAATAGCTCGAGATCAACACACAGACCCAGTGGACACCAACA 960
QY 4516 agctcaaacccagtcaccaagagaagcttcaggaagtgagcttagaagaagaattcttc 4575
D 959 AGCTCCAAGCCACAGTCCAAGAGAGAGTCTCAGGAAAGTGGACGTAGAAGAAGAATTTCTTC 900
QY 4576 gcaactcaggaaaacgaacacacatacagcaggcaagccatgcacacacccccaaacagcagta 4635

Db	899	GCACCTCAGGAAACGAACACCATCAGCAGCAAGCAATGCACACACCCAAACCCAGCAGTA	840
Qy	4636	agtgtgagaaaaacatctacgcatttatgggaactccagtcagagaaactgacctgaca	4695
Db	839	AGTGTGTAGAAAAACATCTACGCATTTATGGGAATCTCCAGTCAGAAACTGGACCTGACA	780
Qy	4696	gagaacttaactggcagcaagagacggtctacaaaactcctaaggaaagcccgagctcta	4755
Db	779	GAGAATCTTAACCTGGCAGCAAGAGACGGCTACAAAATCCTTAAGGAAAGCCCGAGCTCTA	720
Qy	4756	gaagacctggctggttttaagagagctttccagacacagaggtccacactgaggaataaatg	4815
Db	719	GAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACACAGAGGTTCACACTCAGGAATCAATG	660
Qy	4816	actaacgataaaactgcacaaagtacgtgcgaatcttcacacccagactagacaaaaac	4875
Db	659	ACTAACGATAAACTGCCAAGTAGCCTTGCAAACTTTTACAAACACAGACCCAGACAAAAC	600
Qy	4876	ccagcaagctccaaagcgagctcaagacatccctcctggggaagtggcggtgaaagagag	4935
Db	599	CCAGCAAGCTCAAGCGACGGCTCAAGACATCCCTGGGGAAAGTGGCGCTGAAGAAGAG	540
Qy	4936	ctcctagcagttggcaagctcaacagacatacaggagagactacaacacacacacagag	4995
Db	539	CTCCTTAGCAGTTGGCAAGCTTCACACAGACATCAGAGAGACTTACACACACACACAGAG	480
Qy	4996	ccaacagagatggttaagagcatgaagcatttatggagctctccaaagcagatcttagac	5055
Db	479	CCAACAGAGATGTTAAGACGATGAAGCATTTATGGAGTCTCCAAACAGACATCTTAGAC	420
Qy	5056	tcagcagcaagtctaactggcagcaagcagcagctgagaactcctaagggaagtctgaa	5115
Db	419	TCAGCAGCAAGTCTTAACCTGGCAGCAAGAGCGAGCTGAGAACTTCCTTAAGGAAAGTCTGAA	360
Qy	5116	gtccctgaagacctggcgggttcagagctcttcagagacacccaagtcacactaagaaa	5175
Db	359	GTCCCTGAAGACCTGGCGGGCTTCATCGAGCTCTTCCAGACACCAAGTCACACTAAGGAA	300
Qy	5176	tcaatgactaatgaaaaactaccaaagtatcctacagagcttcacagcagacctagt	5235
Db	299	TCAATGACTTACGAAAAAATACCAAAAGTATCTTACAGAGTTCACAGCCACCTAGTG	240
Qy	5236	gacaccccaagaagctccaagccacagcccaagagaagtctcaggaaaagcagacactgaa	5295
Db	239	GACACCCCAACAAGCTCCAAGCCACAGCCCAAGAGAAGTCTCAGGAAAGCAGACACTGAA	180
Qy	5296	gaagaatttttagcatttaggaacaaacgcgcatacagcagcgaagccatgcacacacccc	5355
Db	179	GAAGAAATTTTATAGCATTTTAGGAAACAAAACGCCATCAGCAGGCAAAAGCCATGCACACCCC	120
Qy	5356	aaaccagcagtaggtgagagaaaagacatcaacacgctttttgggaactccagtcagaaa	5415
Db	119	AAACCCAGCAGTAGGTGAGAGAGAAACATCAACACGTTTTTGGGAACCTCCAGTCGACGAAA	60
Qy	5416	ctggaccagccaggaaatttacctggcagcaatagacggctacaaactcgtaaggaaaa	5474
Db	59	CTGGACCCAGCAGGAAATTTTACCTGGCAGCAATAGACGGCTACAAACTCGTAAAGCAAAA	1

RESULT	8				
LOCUS	AX147668	733 bp	DNA	linear	PAT 08-JUN-2001
DEFINITION	Sequence 1 from Patent W00136629.				
ACCESSION	AX147668				
VERSION	AX147668.1	GI:14346724			
KEYWORDS	synthetic construct.				
SOURCE	synthetic construct.				
ORGANISM	artificial sequence.				
REFERENCE	1 (bases 1 to 733)				
AUTHORS	Gerdes,J., Scholzen,T. and Wohlenberg,C.				
TITLE	Transfer compounds, the production and the use thereof				

JOURNAL	Patent: WO 0136629-A 1 25-MAY-2001;				
FEATURES	Forschungszentrum Borstel (DE)				
source	1. 733				
	/organism="synthetic construct"				
	/db_xref="taxon:33630"				
	/note=" Kon21-DNA-Insert, enthaltend einen Teilbereich der				
	fur das Ki-67 Protein kodierenden Sequenz und				
	Restriktionsstellen"				
BASE COUNT	271 a	144 c	199 g	119 t	
ORIGIN					
Query Match	6.4%; Score 629; DB 6; Length 733;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 629; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	9108	gccaagagcaaatcatccgaaccccggtgctcatcatgaagagaagtgttgagactctgc	9167		
Db	62	GGCAAGAGCAAAATCATCCGAACCCGTGCTCATCATGAAGAGAAAGTTTGAGGACTTCTGC	121		
Qy	9168	aaaaagaattgaacctgcggaagagctgaacagcaacgacatgaaacccaacaaagaga	9227		
Db	122	AAAAAGAAATTGAACCTTCGGGAAGAGCTGAACAGCAACGACATGAAACCAACAAAGAGA	181		
Qy	9228	acacaaattacaagactcgtccctgaaataaagggaatatccctcgcgtccagagccca	9287		
Db	182	ACACAAATTTACAAGACTCGTGTCCTTGAAATAAGGGAATATCCCTCGCTCCAGACGCCA	241		
Qy	9288	agataagactcaggcagaacacagaataaactagtgctctttgtattatgcgagaagaataga	9347		
Db	242	AGATAAGACTGAGGCAGACAACAGCAATAACTTGAGGTCTTTGTATTATGACGAGAAGATA	301		
Qy	9348	aataacagaaatgaaagagcccatgaagacctccagagatgacattcagaatcc	9407		
Db	302	AATAACAGAAATGAAAGAACCCCATGAAGACCTCCCCAGAGATGGACATTTCAGAATCC	361		
Qy	9408	agatgatgagcccggaaccccatacctagagacaaagtcaactgagaaacaaaggctgt	9467		
Db	362	AGATGATGAGCCCGGAACCCCATACCTAGAGACAAAGTCACTGAGACAAAGGTTGCTT	421		
Qy	9468	gaggctgctagacagaatgagagctcccagcctaagtggtgcagagagagcggaggca	9527		
Db	422	GAGGTCTGTGACAGAAATGAGAGCTCCAGGCTTAAGGTGGCAGAGAGGAGCGAGGGCA	481		
Qy	9528	gaagagtcgaaggttctcatgcagaatcagaaagggaagagagagcaggaattcaga	9587		
Db	482	GAAGAGTCGAAGGTTCTCATCGAANTCAGAAAGGGAAGGAGAGAGGAAATTCAGA	541		
Qy	9588	ctccatgtccttgagatcaagaaagacaaaagccagcctgcagcgaagcacctttggagag	9647		
Db	542	CTCCATGTGCTTGAGATCAAGAAAGACAAAAGCCAGCCTGCAGCAAGCACTTTGGAGAG	601		
Qy	9648	caaatctgtgcgagagtaacgcgagtgctcaagaggtgtcagaagaaatcccaagaagcc	9707		
Db	602	CAAAATCTGTGCAGAGAGTAACCGGAGTGTCAAGAGGTGTGCAGAAAAATCCRAAAGGC	661		
Qy	9708	tgagacaatgtgtgtcaagaaaaataa	9736		
Db	662	TGAGGCAATGTGTGTCAAGAAAAATAA	690		

RESULT	9				
LOCUS	AX341136	568 bp	DNA	linear	PAT 10-JAN-2002
DEFINITION	Sequence 1383 from Patent W00196398.				
ACCESSION	AX341136				
VERSION	AX341136.1	GI:18137118			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				

REFERENCE 1 (sites)
AUTHORS Jiang,Y., Harlocker,S.L. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0196388-A 1383 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source l..368
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 201 a 144 c 130 g 90 t 3 others
ORIGIN

Query Match 4.7%; Score 456; DB 6; Length 568;
Best Local Similarity 99.8%; Pred. No. 1.4e-244;
Matches 506; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7243 ggaattaccctggcagcagaagacagccacagactccttaaggaaaggctgaggctcta 7302
Db 1 GGAATTTTACCTGGCAGCAAGACAGACGCCACAGACTCCTTAAGGAAAAGGCTGAGGCTCTA 60

QY 7303 gaggacctggttgcttcaaaagactcttccagacaccagggtcacactagagaatcaatg 7362
Db 61 GAGGACTGGTGGCTTCAAAAGAACHTCTCCAGACACCAAGGTCCACACTGAGGAATCAATG 120

QY 7363 actgatgacaaaatacacagaagtatcctgtaaatctccacagccagagtcattcaaaaacc 7422
Db 121 ACTGATGACAAAATCACAGAAGTATCCTGTAAATCTCCACAGCCAGGGTCATTCAAAACC 180

QY 7423 tcaagaagctccaagcaaggctcaagatatacccttggtgaaagtggacatgaagaagag 7482
Db 181 TCAAGAAGCTCCAAGCAAAAGGCTCAAGATACCCCTGGTGAAGTGGACATGAAAGAAGAG 240

QY 7483 cccttagcagtcagcagctcacagcagcatcaggggagactacgcaaacacacacagag 7542
Db 241 CCCTAGCAGTCAGCAAGCTCACAGCGACATCAGGGGAGACTACGCAAAACACACAGAG 300

QY 7543 ccaacaggagatagtagaagcagcatcaagcgtttaaggagctctccaaagcagatccctggac 7602
Db 301 CCAACAGGAGATAGTAAGAGCATCAAAAGCGTTTAAAGGAGTCTCCAAGCAGATCCTGGAC 360

QY 7603 ccagcagcagtgtaactggttagcagagcagcagcagcagcagcagcagcagcagcagc 7662
Db 361 CCAGCAGCAAGTGTAACTGGTAGCAGGAGCGAGCTGAGAACTCGTAAAGAAAAGGCCCGT 420

QY 7663 gctctagaagacctggttgacttcaaaagagctcttctcagcaccaggtcacactgaagag 7722
Db 421 GCTCTAGAGACCTGGTTGACTTCAAAAGAGCTTCTCTCAGCACCAGGTCAACTGAAGAG 480

QY 7723 tcaatgactattgacaaaaacacaaaa 7749
Db 481 TCAATGACTATTGACAAAAACACAAAA 507

RESULT 10
G33252/c 318 bp DNA linear STS 30-JUN-1997
LOCUS human STS SHGC-52534, sequence tagged site.
DEFINITION G33252
ACCESSION G33252
VERSION G33252.1 GI:2226556
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 318)
Myers,R.M.
Unpublished (1997)
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine

Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: TTTCAGCCAGCCAGCAAGTCTT
Primer B: ACCTTTGGGAAAGGGATA
STS size: 226
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds
Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs derived from W87807 -- Unigene.

FEATURES
Source 1..318
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="10"
STS 92..317
primer_bind 92..111
primer_bind complement(298..317)
BASE COUNT 47 a 69 c 87 g 114 t 1 others
ORIGIN

Query Match 2.5%; Score 245; DB 11; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6368 cacccttgggaaaaaggatatagtgaaagctctcagccctgaagcagctcacacaga 6427
Db 318 CACCTTTGGGAAAAGGGATATAGTGGAAAGAGCTCTCAGCCCTGAAGCAGCTCACACAGA 259

QY 6428 ccacacacacagacaaaagtaccaggagatgagataaaggcatcaacgtgttcaggaaa 6487
Db 258 CCACACACACACACAAAGTACCAGGAGATGAGGATAAAGGCATCAACGTGTTCAGGGAAA 199

QY 6488 ctgcacacacagaactggaccagcagcagcaagtgttaactggttagcagagcgccaaagaa 6547
Db 198 CTGCAAAACAGAAACTGGACCCAGCAGCAAGTCTAACTGTAGCAAGAGCAGCCCAAGAA 139

QY 6548 ctccctaaggaaaaagcccaaccccttagaagacttgctggttgaagagctcttcaga 6607
Db 138 CTCTTAAGGGAAGCCCAACCCCTTAGAAGACTTGGCTGGCTTGAAGAGCTCTTCCAGA 79

QY 6608 cacca 6612
Db 78 CACCA 74

RESULT 11
G03247/c 332 bp DNA linear STS 19-OCT-1995
LOCUS human STS WI-1011, sequence tagged site.
DEFINITION G03247
ACCESSION G03247
VERSION G03247.1 GI:720205

KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human genome wide STSs created from sheared whole human DNA.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 332)
JOURNAL Hudson, T.
REFERENCE Whitehead Institute/MIT Center for Genome Research; Random Genome
AUTHORS Wide STSs
TITLE 2 (bases 1 to 332)
JOURNAL Hudson, T.
REFERENCE Whitehead Institute/MIT Center for Genome Research; Physically
AUTHORS Mapped STSs
TITLE Unpublished
JOURNAL GDB_DSEG: D10S1249
COMMENT Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: CTTCTGACCTGTTTGAGT
Primer B: CTTGAGTGCACAGATGACG
STS size: 201
PCR Profile:
Presoak:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 mM
Tag Polymerase: 0.025 units/ul
Total Vol: 20 ul
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCL: 10 mM
PH: 9.3.
Location/Qualifiers
1..332
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="791_D_7"
65..265
65..84
primer_bind
primer_bind
complement(246...265)
BASE COUNT 76 a 74 c 67 g 115 t
ORIGIN
Query Match 2.0%; Score 199; DB 11; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.4e-99;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2416 ggaggaatgtgttttcagtagcacagaaatgcagcaaacagccatctgataaaatgctct 2475
|||||
DB 279 GAGGAAATGTGTTCTTCAGTCACAGATGCAGCAAAACAGCCATCTGATAAATGCTCT 220
QY 2476 gcaagccctcccttaagacgcagtgattagagaaatggaaacgtagcaaaaacgccc 2535
|||||
DB 219 GCAAGCCCTCCCTTAAGACGGCAGTGATTAGAGAAATGGAACAGTACCAAAACGCC 160
QY 2536 aggaacacctaacaatgacttctctgagacaaaacttcagatactgagacagagcct 2595
|||||

Db 159 AGGAACACCTACAAATGACTTCTCTGGAGACAAAAAAGTTCAGATACTGAGACAGAGCCT 100
QY 2596 tcaaaaacagtatccactg 2614
|||||
Db 99 TCAAAAACAGTATCCACTG 81
RESULT 12
LOCUS AX336882/c 409 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 7391 from Patent WO0194629.
ACCESSION AX336882
VERSION AX336882.1 GI:18127601
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 7391 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
Source
1..409
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 98 a 90 c 84 g 137 t
ORIGIN
Query Match 2.0%; Score 198; DB 6; Length 409;
Best Local Similarity 100.0%; Pred. No. 5.1e-99;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 493 gacgtaccgcagatgactcaaaagacagtggttcacaggaacaaactaatgtctatcc 552
|||||
Db 198 GACAGTACCGCAGATGACTCAAAAGACAGCTGTTGTCAGGGACACACTAATGTTCAATCC 139
QY 553 tcagaacatgctggacgtaagtgcagaaatgcagctgataccattctctgggattttaa 612
|||||
Db 138 TCAGAACATGCTGGAGCGTAATGCGAAATGCAGCTGATCCCATTTCTGGGATTTTAA 79
QY 613 gaaattccagcgttaaatagtagccgttatgagaaatgagagctcttcccaactaca 672
|||||
Db 78 GAAATTCAGCGTTAAATAGTAGCCGTTATGGAGAAATGAAGTCTGTTCCCACTACA 19
QY 673 caatgtcttgacaatagc 690
|||||
Db 18 CAATGTCTTGACAAATAGC 1
RESULT 13
LOCUS AX072812 158 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 3284 from Patent WO0102568.
ACCESSION AX072812
VERSION AX072812.1 GI:12583165
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 158)
AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
Lanson, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M.,
Labat, I., Lesnikowicz, D., Kita, D., Garcia, V. and Strache-Crain, B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 3284 11-JAN-2001;
CHIRON CORPORATION (US) ; HYSEQ, INC. (US)

```
FEATURES
Source      Location/Qualifiers
misc_feature 1..158
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
BASE COUNT  54 a 40 c 34 g 27 t 3 others
ORIGIN
Query Match      1..1%: Score 110; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.3e-49;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4699 aacttaactggcagacgagcgtctacaaactctctaaagaaagcccgaggtctctagaa 4758
|||||
Db 11 AACTTAAGTGGCAGCAGACGCGCTACAACTCTTAAGGAAAGGCCAGGCTCTAGAA 70
|||||

QY 4759 gacctggctggctttaagagctcttcacagacagcggtcacacactgagga 4808
|||||
Db 71 GACCTGGCTGGCTTTAAAGAGCTCTTCACAGACACGAGGTCACACTGAGGA 120
|||||

RESULT 14
HSK167
LOCUS      HSK167
DEFINITION H. sapiens DNA for Ki-67 antigen 5'-region (exon 1 & 2).
ACCESSION X94762
VERSION    X94762.1 GI:1944550
KEYWORDS   Ki-67 gene; monoclonal antibody.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 14041)
AUTHORS    Gerdes,J.
TITLE      Direct Submission
JOURNAL    Submitted (08-JAN-1996) J. Gerdes, Molecular Immunology,
            Forschungszentrum Borstel, Parkallee 22, D- 23845, Borstel, FRG
REFERENCE  2 (bases 1 to 14041)
AUTHORS    Gerdes,J.
TITLE      Sequence of the human Ki-67 protein gene 5' and promoter region
JOURNAL    Unpublished
COMMENT    On Apr 18, 1997 this sequence version replaced gi:1869800.
FEATURES
Source      Location/Qualifiers
misc_feature 1..14041
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="10"
             /map="q25"
             /clone="EMBL3"
             /cell_type="HeLa S3"
             /clone_lib="HeLa S3"
mRNA        join(11797..11903,12245..>12430)
            /gene="Ki-67"
            /number=1
exon        11797..11903
            /gene="Ki-67"
            /number=1
gene        11797..12430
            /gene="Ki-67"
            /number=2
intron      11904..12244
            /gene="Ki-67"
            /number=1
exon        12245..>12430
            /gene="Ki-67"
            /number=2
CDS         12335..12430
            /codon_start=1
            /product="monoclonal antibody Ki-67"
            /protein_id="CAA64388.1"
            /db_xref="GI:1869801"
            /db_xref="SWISS-PROT:P46013"

BASE COUNT  3477 a 3189 c 3327 g 4048 t
ORIGIN
/translation="MMPTRLRLVTIKRSGVDGPHPLSLSTCLFGR"

Query Match      1.0%; Score 93; DB 9; Length 14041;
Best Local Similarity 100.0%; Pred. No. 8.6e-40;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgccccagagcgcctgttactacaaagagcggggtcgagctccccacttt 60
|||||
Db 12335 ATGTGGCCCCAGAGACGCCCTGGTTACTATCAAAAGGAGCGGGTCTGACGCTCCCATTT 12394
|||||

QY 61 ccctgagcctcagcacctgctgtttgttggagg 93
|||||
Db 12395 CCCTGAGCCTCAGCACCTGCTGTGTTTGGGAGG 12427
|||||

RESULT 15
HS71L16/c
LOCUS      HS71L16
DEFINITION Human DNA sequence from clone 71L16 on chromosome Xp11. Contains a
            probable Zinc Finger protein (pseudo)gene, an unknown putative
            gene, a pseudogene with high similarity to part of antigen KI-67, a
            putative Chondroitin 6-Sulfotransferase LIKE gene and a KIAA0267
            LIKE putative Na(+)/H(+) exchanger protein gene. Contains a
            predicted CpG island, ESTs, STSs and GSSs and genomic markers
            DXS1003 and DXS1055, complete sequence.
ACCESSION  AL022165
VERSION     AL022165.1 GI:3281985
KEYWORDS    HTG; antigen KI-67; CpG island; DXS1003; DXS1055; KIAA0267;
            Sodium/Hydrogen exchanger; Zinc Finger.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 176433)
AUTHORS    Grafham,D.
TITLE      Direct Submission
JOURNAL    Submitted (07-JUL-1998) E-mail enquires: humquery@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT    On Jul 1, 1998 this sequence version replaced gi:3250841.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above. This sequence is the
            entire insert of clone 71L16. This sequence has been finished
            according to sequence map criteria as follows. An attempt is made
            to resolve all sequencing problems, such as compressions and
            repeats, but not necessarily within known annotated human repeat
            sequence elements (e.g. Alu). Where the sequence is ambiguous,
            there is an annotation using the 'unsure' feature key.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome X, constructed by the Sanger Centre Chromosome X
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/ChrX
            71L16 is from the library RPC11 constructed at the Roswell Park
            Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/ VECTOR: PCYPAC2.
FEATURES
Source      Location/Qualifiers
misc_feature 1..176433
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="X"
             /map="p11"
             /clone="RP1-71L16"
             /clone_lib="RPC1-1"
repeat_region 1..259
            /note="Alu repeat: matches 59. .301 of consensus"
repeat_region 271..558
            /note="AluX repeat: matches 1. .289 of consensus"
```



```
repeat_region 675..974
/note="AluX repeat: matches 1..300 of consensus"
repeat_region 1077..1465
/note="L1M2 repeat: matches 375..772 of consensus"
repeat_region 1469..1767
/note="AluX repeat: matches 1..302 of consensus"
repeat_region 1785..1981
/note="L1M1 repeat: matches 772..963 of consensus"
repeat_region 2254..2572
/note="AluX repeat: matches 1..301 of consensus"
repeat_region 2649..2931
/note="AluJb repeat: matches 6..299 of consensus"
repeat_region 2942..3254
/note="AluX repeat: matches 1..299 of consensus"
repeat_region 3299..3478
/note="L1M10 repeat: matches 40..214 of consensus"
repeat_region 3392..3490
/note="L1M1 repeat: matches 128..225 of consensus"
repeat_region 3551..3730
/note="L1M2 repeat: matches 326..516 of consensus"
repeat_region 3733..4030
/note="AluJb repeat: matches 1..300 of consensus"
repeat_region complement(4678..4977)
/note="AluX repeat: matches 300..1 of consensus"
repeat_region 5111..5409
/note="L1 repeat: matches 2..301 of consensus"
repeat_region 5414..5745
/note="L1 repeat: matches 3481..3826 of consensus"
repeat_region 5755..6053
/note="AluX repeat: matches 1..303 of consensus"
repeat_region 6083..6373
/note="AluJb repeat: matches 6..297 of consensus"
repeat_region 6404..6651
/note="L1 repeat: matches 3880..4128 of consensus"
repeat_region 6670..6988
/note="AluS repeat: matches 1..303 of consensus"
repeat_region 6990..7615
/note="L1 repeat: matches 4148..4780 of consensus"
repeat_region complement(7634..7926)
/note="AluJb repeat: matches 295..1 of consensus"
repeat_region 8335..8627
/note="AluS repeat: matches 2..299 of consensus"
repeat_region 8631..8923
/note="L1M1 repeat: matches 41..346 of consensus"
complement(join(<9684..9794,10132..10258,23422..23465,
25850..25960,26720..26768))
/gene="dJ7L1L16.1"
/note="match: cDNA D31763
match: ESTs AA083688 N41340
match: proteins P51523 Q99676 P21506 Q06730 Q61967 Q62977
Q13401 Q14628 Q14898 Q03938 Q13359 Q13360 Q15636 Q15928
Q43337 P52738 Q14588 P70591 Q61116 Q14913 Q62788 P52737
O15322 Q61491 Q64247 Q62515 P70405
could be a pseudogene"
/product="dJ7L1L16.1 (probable Zinc Finger protein)"
/evidence=not_experimental
complement(join(9684..9794,10132..10258,23422..23465,
25850..25960,26720..26768))
/gene="dJ7L1L16.1"
complement(10332..10516)
/note="MIR repeat: matches 242..23 of consensus"
10738..11036
/note="AluS repeat: matches 1..299 of consensus"
complement(11993..12280)
/note="MER21B repeat: matches 790..499 of consensus"
12284..12453
/note="AluJb repeat: matches 135..302 of consensus"
complement(12471..12954)
/note="MER21B repeat: matches 490..6 of consensus"
13147..13181
/note="L1P45 repeat: matches 856..890 of consensus"
13234..13536
/note="AluS repeat: matches 1..302 of consensus"

repeat_region 13925..14083
/note="LTR8 repeat: matches 1..146 of consensus"
repeat_region 14101..14326
/note="MER30 repeat: matches 5..229 of consensus"
repeat_region 14632..14815
/note="LTR8 repeat: matches 504..691 of consensus"
repeat_region 14830..15029
/note="AluX repeat: matches 2..201 of consensus"
repeat_region 15180..15219
/note="20 copies 2 mer ta 83% conserved"
complement(15257..15621)
/note="THE1a repeat: matches 364..2 of consensus"
complement(15709..16009)
/note="AluJb repeat: matches 283..1 of consensus"
complement(16088..16389)
/note="AluS repeat: matches 303..1 of consensus"
complement(16494..16794)
/note="AluJb repeat: matches 301..1 of consensus"
17527..17657
/note="AluJb repeat: matches 1..132 of consensus"
17659..17955
/note="AluS repeat: matches 1..299 of consensus"
17957..18125
/note="AluX repeat: matches 134..302 of consensus"
18338..18624
/note="AluX repeat: matches 1..292 of consensus"
19156..19451
/note="AluX repeat: matches 1..296 of consensus"
19959..20250
/note="AluJb repeat: matches 1..292 of consensus"
20253..20553
/note="AluJb repeat: matches 2..302 of consensus"
20832..20929
/note="L1M3 repeat: matches 458..556 of consensus"
20953..21164
/note="MER20 repeat: matches 1..217 of consensus"
21312..21611
/note="AluY repeat: matches 1..300 of consensus"
22131..22428
/note="AluS repeat: matches 1..298 of consensus"
22429..22730
/note="AluS repeat: matches 1..303 of consensus"
23798..24080
/note="AluJb repeat: matches 21..302 of consensus"
complement(24297..24602)
/note="AluS repeat: matches 303..1 of consensus"
complement(26186..26312)
/note="MIR2 repeat: matches 145..1 of consensus"
26366..26464
/note="MIR repeat: matches 74..177 of consensus"
join(<26877..26976,27815..>28001)
/gene="dJ7L1L16.2"
/note="match: EST AA322753"
/product="dJ7L1L16.2 (putative protein)"
/evidence=not_experimental
join(26877..26976,27815..28001)
/gene="dJ7L1L16.2"
complement(28826..29128)
/note="AluS repeat: matches 297..1 of consensus"
29139..29824
/note="match: ESTs AA487141 AA996220 P20986 AA492440
AA507555 AA486818
match: protein Q14585"
complement(29207..29456)
/note="AluJb repeat: matches 245..1 of consensus"
complement(29916..29985)
/note="L1M9 repeat: matches 983..910 of consensus"
30031..30328
/note="AluX repeat: matches 1..297 of consensus"
complement(30332..30521)
/note="L1M8 repeat: matches 777..600 of consensus"
30810..31099
/note="AluY repeat: matches 1..294 of consensus"
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2002, 17:45:00 ; Search time 1314.37 Seconds
(without alignments)
12756.976 Million cell updates/sec

Title: US-09-700-906A-1_COPY_197_9962
Perfect score: 9766
Sequence: 1 atgtgcccacgagacgct.....tcatagggacagtaagata 9766

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : N_Geneseq_032802.*
- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
 - 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
 - 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
 - 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
 - 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
 - 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
 - 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
 - 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
 - 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
 - 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
 - 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
 - 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
 - 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
 - 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
 - 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
 - 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
 - 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
 - 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
 - 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
 - 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
 - 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
 - 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
 - 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
 - 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match	Length	ID		
1	9766	100.0	12493	21	AAZ43872	Human cell cycle p
2	9232	94.5	12515	21	AAA64658	cDNA sequence enco
3	8286	84.8	11435	21	AAA64659	cDNA sequence enco
4	629	6.4	733	22	AAH20873	Human Ki-67 protei
5	527	5.4	904	21	AAA64645	Partial sequence M
6	389	4.0	561	21	AAZ80374	Human colon cancer
7	370	3.8	850	21	AAA64643	Partial sequence M
8	344	3.5	861	21	AAA64639	Partial sequence M
9	336	3.4	950	21	AAA64644	Partial sequence M

10	299	3.1	1094	23	AAS69164	DNA encoding novel
c 11	134	1.4	201	19	AAI10717	Human biallelic po
12	110	1.1	158	22	AAF67522	Novel human polynu
13	78	0.8	219	16	AAT20857	Human gene signatu
14	64	0.7	1126	23	AAS69163	DNA encoding novel
15	51	0.5	51	22	AAI27414	Human SNP oligonuc
16	51	0.5	51	22	AAI32696	Human SNP oligonuc
17	51	0.5	51	22	AAI32697	Human SNP oligonuc
18	51	0.5	51	22	AAI32698	Human SNP oligonuc
19	51	0.5	51	22	AAI33289	Human SNP oligonuc
20	51	0.5	51	22	AAI33290	Human SNP oligonuc
21	51	0.5	51	22	AAH79977	Human DNA containi
22	27	0.3	51	22	AAI34120	Human SNP oligonuc
23	26	0.3	53	21	AAA64666	PCR primer used to
24	25	0.3	51	22	AAI27415	Human SNP oligonuc
25	25	0.3	51	22	AAI33291	Human SNP oligonuc
26	24	0.2	24	19	AAI30923	Human biallelic po
c 27	24	0.2	24	19	AAI09349	Human biallelic po
c 28	24	0.2	36	21	AAA64667	PCR primer used to
c 29	23	0.2	23	21	AAZ43869	Human cell cycle p
30	23	0.2	23	21	AAZ43870	Human cell cycle p
31	21	0.2	343	21	AAC29850	Human secreted pro
32	21	0.2	417	22	AAI37373	Human musculoskele
33	21	0.2	417	22	AAI37375	Human musculoskele
34	21	0.2	417	22	AAI37378	Human musculoskele
35	21	0.2	457	22	AAI81035	Human polynucleoti
c 36	21	0.2	489	22	AAI93620	Human polynucleoti
c 37	21	0.2	598	22	AAI60943	Human polynucleoti
c 38	21	0.2	2555	22	AAI59157	Human polynucleoti
c 39	21	0.2	8956	23	ABLI2828	Human polynucleoti
c 40	21	0.2	14388	23	ABLI05670	Drosophila melanog
41	21	0.2	23071	22	AAS26699	Human genomic DNA
c 42	20	0.2	340	17	AAT43951	Sequence flanking
43	20	0.2	422	22	AAF64763	Novel human polynu
44	20	0.2	469	22	ABA43593	Human breast cell
45	20	0.2	469	22	ABA54045	Human foetal liver

ALIGNMENTS

RESULT 1	
AAZ43872	
ID	AAZ43872 standard; DNA; 12493 BP.
XX	
AC	AAZ43872;
DT	10-MAR-2000 (first entry)
XX	
DE	Human cell cycle protein Ki-67 DNA.
XX	
KW	Cell cycle protein; Ki-67; therapy; cell proliferation; allergy;
KW	tumor treatment; autoimmune disease; scar formation; inflammation;
KW	rheumatic disease; transplantation; ds.
XX	
OS	Homo sapiens.
XX	
EH	Key
FT	CDS
FT	Location/Qualifiers
FT	197..9967
FT	/*tag= a
XX	
XX	/product= "Ki-67"
PN	DE19822954-A1.
XX	
PD	25-NOV-1999.
XX	
PF	22-MAY-1998; 98DE-1022954.
XX	
PR	22-MAY-1998; 98DE-1022954.
XX	
PA	(BORS-) FORSCHUNGSZENTRUM BORSTEL ZENT MEDIZIN.
XX	
PI	Flad H, Gerdes J, Boehle A, Deinert I;

XX WPI; 2000-039964/04.
DR P-PSDB; AAY50976.
XX
XX K1-67 gene antisense oligonucleotide -
XX
XX Disclosure; Page 5-19; 36pp; German.
XX
CC This invention describes a novel oligoribo- or oligodeoxyribonucleotide,
CC characterized in that, it hybridizes to mRNA that encodes protein Ki-67
CC at a physiologically acceptable salt concentration. The oligoribo- or
CC oligodeoxyribonucleotide which is complementary to Ki-67, a protein
CC active at all stages of the cell cycle except G₀, is useful for therapy
CC of illnesses with increased cell proliferation and particularly for
CC treatment of tumors, autoimmune diseases, scar formation, inflammation,
CC allergy, rheumatic diseases and defence against transplantation. This
CC sequence encodes the human cell cycle protein Ki-67 which is described in
CC the method of the invention.
XX
SQ Sequence 12493 BP; 4143 A; 3048 C; 2929 G; 2373 T; 0 other;

Query Match 100.0%; Score 9766; DB 21; Length 12493;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9766; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgcccacgagcgcctgttactatcaaaagagcgggtcgacggtcccccacttt 60
Db 197 atgtgcccacgagcgcctgttactatcaaaagagcgggtcgacggtcccccacttt 256
QY 61 ccctcgagcctcgacccctgctgtttggaaggggtattgaattgacatccgtatccag 120
Db 257 ccctcgagcctcgacccctgctgtttggaaggggtattgaattgacatccgtatccag 316
QY 121 ctccctgtgttcacaaacacattgcaagttgaattccatcgacgagggcaatatita 180
Db 317 ctccctgtgtgttcacaaacacattgcaagttgaattccatcgacgagggcaatatita 376
QY 181 cataatttcagttccacaaatccacaaagtaattcccaagaaatacgtgaacaggagcca 240
Db 377 cataatttcagttccacaaatccacaaagtaattcccaagaaatacgtgaacaggagcca 436
QY 241 cggctaaaaacattggaatgtaataactattattgactgttccttcagggtatgaaaaatgaa 300
Db 437 cggctaaaaacattggaatgtaataactattattgactgttccttcagggtatgaaaaatgaa 496
QY 301 agttctcagaatggaaggaaactcaactgaattcccaagaaatacgtgaacaggagcca 360
Db 497 agttctcagaatggaaggaaactcaactgaattcccaagaaatacgtgaacaggagcca 556
QY 361 gcacgtcgtgtccaagatcagcttctctcttcgacccctgatgagaagctcaagattcc 420
Db 557 gcacgtcgtgtccaagatcagcttctctcttcgacccctgatgagaagctcaagattcc 616
QY 421 aaggcctattcaaaaatcactgaaggaaagtttcaggaaatccctcaggtacatatcaag 480
Db 617 aaggcctattcaaaaatcactgaaggaaagtttcaggaaatccctcaggtacatatcaag 676
QY 481 aatgtcaagaagacagataccgcagatgactcaaaagacagtggttgcagggaacaact 540
Db 677 aatgtcaagaagacagataccgcagatgactcaaaagacagtggttgcagggaacaact 736
QY 541 aatgttcattctcagaacatgctggacgttaattggaagaaatgcagctgacccattct 600
Db 737 aatgttcattctcagaacatgctggacgttaattggaagaaatgcagctgacccattct 796
QY 601 ggggattttaaaagaatttccagcgttaaatagtgcggttatggagaattgaagctt 660
Db 797 ggggattttaaaagaatttccagcgttaaatagtgcggttatggagaattgaagctt 856
QY 661 gttcccaactacacaaatgtcttgacaaatagcaaaaaaaatgaattccctcttggagagctt 720
Db 857 gttcccaactacacaaatgtcttgacaaatagcaaaaaaaatgaattccctcttggagagctt 916

QY 721 tatgagtcagtgaaagaaagagttggatgtataaatcacaaaaagaaaaatgtcctacagtat 780
Db 917 tatgagtcagtgaaagaaagagttggatgtataaatcacaaaaagaaaaatgtcctacagtat 976
QY 781 tgtagaaaaatctggtattacaaaactgattacgcacacagagagaaagagtgctgagtgtta 840
Db 977 tgtagaaaaatctggtattacaaaactgattacgcacacagagagaaagagtgctgagtgtta 1036
QY 841 caggaggagaccacaactgttgcctcgcgttaagtcaagaccacaaaatcgtgtggagcgcgc 900
Db 1037 cagggggagaccacaactgttgcctcgcgttaagtcaagaccacaaaatcgtgtggagcgcgc 1096
QY 901 cagcgtgtgacagacgctgcttcacctgaaacaagagcttgaccagacaaggggaagggga 960
Db 1097 cagcgtgtgacagacgctgcttcacctgaaacaagagcttgaccagacaaggggaagggga 1156
QY 961 agagacgtggagctctgttcagactccacagcaagcgtgtgggcccagcgttctctctctat 1020
Db 1157 agagacgtggagctctgttcagactccacagcaagcgtgtgggcccagcgttctctctctat 1216
QY 1021 gagccgcctaaaaatgaagaccctctgataaatttcacagcaacaaaaatctccacaaaaa 1080
Db 1217 gagccgcctaaaaatgaagaccctctgataaatttcacagcaacaaaaatctccacaaaaa 1276
QY 1081 cataagaacaaaaagacctgtatactactggttagaagagaaatctgtgaatctgggttaaaagt 1140
Db 1277 cataagaacaaaaagacctgtatactactggttagaagagaaatctgtgaatctgggttaaaagt 1336
QY 1141 gaaggctcaagcgtggtgataaaaaatcttactccacaggaagctttcaactagaaatcga 1200
Db 1337 gaaggctcaagcgtggtgataaaaaatcttactccacaggaagctttcaactagaaatcga 1396
QY 1201 acaccagctaaaagtgaagatgcagctgactctgcacataagccacagcaaaaaatctctctcc 1360
Db 1397 acaccagctaaaagtgaagatgcagctgactctgcacataagccacagcaaaaaatctctctcc 1456
QY 1261 aaaaaccagaggaagtattctctacagatgtggaagttctgcctacggaactgaaattcac 1320
Db 1457 aaaaaccagaggaagtattctctacagatgtggaagttctgcctacggaactgaaattcac 1516
QY 1321 aatgagccatttttaactctgtggtcactcaagtgtgagaggaagatccaaaaaggattcc 1380
Db 1517 aatgagccatttttaactctgtggtcactcaagtgtgagaggaagatccaaaaaggattcc 1576
QY 1381 ctacgacagcctgagaaattgggcactacagctgcgacagatgtgctctgggttaccctggt 1440
Db 1577 ctacgacagcctgagaaattgggcactacagctgcgacagatgtgctctgggttaccctggt 1636
QY 1441 cttagttcagttgatatcaacaactttggtgattccatttaattgagagtgagggaatacct 1500
Db 1637 cttagttcagttgatatcaacaactttggtgattccatttaattgagagtgagggaatacct 1696
QY 1501 ttgaaaaagaggcgtgtgctcttgggtggccactcaagaccctgaactatttggatgaaaaac 1560
Db 1697 ttgaaaaagaggcgtgtgctcttgggtggccactcaagaccctgaactatttggatgaaaaac 1756
QY 1561 ttgctctcctaaatcacgctctcaaaaggggagagcccccacacaaaagaaagttctctggta 1620
Db 1757 ttgctctcctaaatcacgctctcaaaaggggagagcccccacacaaaagaaagttctctggta 1816
QY 1621 atgcacactccacctgctctgaagaaaaatcatcaagggaacagcctcaaccatcaggaaaaa 1680
Db 1817 atgcacactccacctgctctgaagaaaaatcatcaagggaacagcctcaaccatcaggaaaaa 1876
QY 1681 caagagtcagggttcagaaatccatgtggaagtgaaaggcacaagcttggttataaagccct 1740
Db 1877 caagagtcagggttcagaaatccatgtggaagtgaaaggcacaagcttggttataaagccct 1936
QY 1741 ccagctccctagtccttaggaaaaactccagttgccagtgatcaacgcgtaggctctctcaaa 1800
Db 1937 ccagctccctagtccttaggaaaaactccagttgccagtgatcaacgcgtaggctctctcaaa 1996

Qy	1801	acagccctctctccagcagcaaatctcagacagaggttctctaagagagagagagagaa	1866
Db	1997	acagccctctctccagcagcaaatctcagacagaggttctctaagagagagagagaa	2056
Qy	1861	gtgcaacctgccttcaaaagagaggtctctatcagcgaagtcacatgatatttcag	1920
Db	2057	gtgcaacctgccttcaaaagagaggtctctatcagcgaagtcacatgatatttcag	2116
Qy	1921	atgatgttccaaaagaagaagtgtgcttcggaagcaaatctcgattgttgcaaaatca	1980
Db	2117	atgatgttccaaaagaagaagtgtgcttcggaagcaaatctcgattgttgcaaaatca	2176
Qy	1981	tgggcagatgtagtataaaccttgggtcaaaacaaacacaaactaaagtctataaacatcgt	2040
Db	2177	tgggcagatgtagtataaaccttgggtcaaaacaaacacaaactaaagtctataaacatcgt	2236
Qy	2041	cctcaaggtcctaagaacaaaggcaagaagacctgctactcaagaagacctgtgggc	2100
Db	2237	cctcaaggtcctaagaacaaaggcaagaagacctgctactcaagaagacctgtgggc	2296
Qy	2101	gaagttccagtcacaaatttagtagacggccacgcaaaactctctgtccataataatagg	2160
Db	2297	gaagttccagtcacaaatttagtagacggccacgcaaaactctctgtccataataatagg	2356
Qy	2161	aaagtcctactgaaaaagtcacatgtcctgctcgacctacagatgctccaacaacttc	2220
Db	2357	aaagtcctactgaaaaagtcacatgtcctgctcgacctacagatgctccaacaacttc	2416
Qy	2221	atttccaaacaaaattggacctttaaggaagatctttcaggaaatagctgaattgttcaag	2280
Db	2417	atttccaaacaaaattggacctttaaggaagatctttcaggaaatagctgaattgttcaag	2476
Qy	2281	acccagtgaaaggagcaaccgcagttgtacaagcacatgtcacatcgcatttccaaattca	2340
Db	2477	acccagtgaaaggagcaaccgcagttgtacaagcacatgtcacatcgcatttccaaattca	2536
Qy	2341	gagaatttcttggaaaacagtttccaaggaactgattcaggaggaagaacctctgctcccc	2400
Db	2537	gagaatttcttggaaaacagtttccaaggaactgattcaggaggaagaacctctgctcccc	2596
Qy	2401	acctcagagagtttggaggaaatgttcttcagtcacagaatcagcaaaacagcca	2460
Db	2597	acctcagagagtttggaggaaatgttcttcagtcacagaatcagcaaaacagcca	2656
Qy	2461	tctgataaatgtctctgaagccctcccttaagacggcagtgattagagaaaaatgaaac	2520
Db	2657	tctgataaatgtctctgaagccctcccttaagacggcagtgattagagaaaaatgaaac	2716
Qy	2521	gtagcaaaacgcccgaggaacacctacaaaatgactctcttgggagacaaaacttcagat	2580
Db	2717	gtagcaaaacgcccgaggaacacctacaaaatgactctcttgggagacaaaacttcagat	2776
Qy	2581	actgagacagagccttcaaaaacagtatccactgtaaacaggttcaggaaaggtctacagag	2640
Db	2777	actgagacagagccttcaaaaacagtatccactgtaaacaggttcaggaaaggtctacagag	2836
Qy	2641	ttcaggaaataacagaaagctacctgtggaagtagagtgaaagaaacaaatcacagaaatt	2700
Db	2837	ttcaggaaataacagaaagctacctgtggaagtagagtgaaagaaacaaatcacagaaatt	2896
Qy	2701	gttgagtgcactctaaaaagaggttcagaaggccaacctactacacaaagagagagga	2760
Db	2897	gttgagtgcactctaaaaagaggttcagaaggccaacctactacacaaagagagagga	2956
Qy	2761	gagatgaaggaaatagaaagacctttgagacataataaggaaaaatattgaaataaagaa	2820
Db	2957	gagatgaaggaaatagaaagacctttgagacataataaggaaaaatattgaaataaagaa	3016
Qy	2821	aacgatgaaaaatgaaagcaatgaagagatcaagaaccttggggcgagaatgtgcacca	2880
Db	3017	aacgatgaaaaatgaaagcaatgaagagatcaagaaccttggggcgagaatgtgcacca	3076
Qy	2881	atgtctgacctgcagacacctcaagaccttgcctgtatcagaaactcatgaaagacacagcca	2940

[illegible]

Db 4157 gacctgacagaaacttaacccggcagcaagagacggccacaaaactcctaaggaaagccc 4216
QY 4021 caggctctgaaagacctgactggctttaagagagctcttccagaccctggctcatactgaa 4080
Db 4217 caggctctgaaagacctgactggctttaagagagctcttccagaccctggctcatactgaa 4276
QY 4081 gaagcagtggctctgctgcaaaactactataaatgcccctgcgaattcttccaccagaaatca 4140
Db 4277 gaagcagtggctctgctgcaaaactactataaatgcccctgcgaattcttccaccagaaatca 4336
QY 4141 gcagacaccccaacacagacaaagaagcagcccaagacacacttttgagaaaagggaagctga 4200
Db 4337 gcagaçcccccacacagacaaagaagcagcccaagacacacttttgagaaaagggaagctga 4396
QY 4201 cagaagagctctcagccctgaagaagctcacacagacatcaggggaaaccacacacaca 4260
Db 4397 cagaagagctctcagccctgaagaagctcacacagacatcaggggaaaccacacacaca 4456
QY 4261 gataaagtacacagaggtgaggataaaagcatcaacgcgttttagggaaaactgcaaaacag 4320
Db 4457 gataaagtacacagaggtgaggataaaagcatcaacgcgttttagggaaaactgcaaaacag 4516
QY 4321 aaactgaccacagcagcaagtgaactggtagcaagagcagcccaaaaactaaggaaaag 4380
Db 4517 aaactgaccacagcagcaagtgaactggttagcaagagcagcccaaaaactaaggaaaag 4576
QY 4381 gccaaacccttagaagacctggctggctggaagagctcttccagacacacagtatgcact 4440
Db 4577 gccaaacccttagaagacctggctggctggaagagctcttccagacacacagtatgcact 4636
QY 4441 gacaagcccacgactcagcagaaactaccaaaatagcctgcagatcacacacagaccaca 4500
Db 4637 gacaagcccacgactcagcagaaactaccaaaatagcctgcagatcacacacagaccaca 4696
QY 4501 gtggacacacaaacagctccaagccacagctccaagagaagctcaggaaaagtggacgtga 4560
Db 4697 gtggacacacaaacagctccaagccacagctccaagagaagctcaggaaaagtggacgtga 4756
QY 4561 gaagaagaattcttcgactcaggaagacgacacccatcagcaggcaagccatgcacaca 4620
Db 4757 gaagaagaattcttcgactcaggaagacgacacccatcagcaggcaagccatgcacaca 4816
QY 4621 ccacaaaccagcagtaagtgctgagaaaaacatctacgcatcttatggaaaactccagtcgag 4680
Db 4817 ccacaaaccagcagtaagtgctgagaaaaacatctacgcatcttatggaaaactccagtcgag 4876
QY 4681 aaactggacctgacagagaacttaactggcagcaagacagcgctacaaaactcctaaggaa 4740
Db 4877 aaactggacctgacagagaacttaactggcagcaagacagcgctacaaaactcctaaggaa 4936
QY 4741 aaggccaggtctagaagacctggctggctttaagagagctcttccagacacagaggtcac 4800
Db 4937 aaggccaggtctagaagacctggctggctttaagagagctcttccagacacagaggtcac 4996
QY 4801 actgaggaatcaatgactaactgataaactgccaagtgccaagtgagcctgcgaattcttcaacaca 4860
Db 4997 actgaggaatcaatgactaactgataaactgccaagtgccaagtgagcctgcgaattcttcaacaca 5056
QY 4861 gacctagacaaaacccagcaagctccaagcgagcgctcaagacatccctgggaaagtga 4920
Db 5057 gacctagacaaaacccagcaagctccaagcgagcgctcaagacatccctgggaaagtga 5116
QY 4921 ggcgtgaaaagaagagctcctagcagttggcaggtcacacagacatcaggagagactaca 4980
Db 5117 ggcgtgaaaagaagagctcctagcagttggcaggtcacacagacatcaggagagactaca 5176
QY 4981 cacacacacagagcccaacagagatggttaagagacatgaaaacatcttatggaggtctcca 5040
Db 5177 cacacacacagagcccaacagagatggttaagagacatgaaaacatcttatggaggtctcca 5236
QY 5041 aagcagatcttagactcagcagcaagcttaactggcagcaagagcgagctgagaactcct 5100
Db 5237 aagcagatcttagactcagcagcaagcttaactggcagcaagagcgagctgagaactcct 5296

QY 5101 aagggaaaagtctgaagtccctgaaagacctggccgggtcttcagagctcttccagacacca 5160
Db 5297 aagggaaaagtctgaagtccctgaaagacctggccgggtcttcagagctcttccagacacca 5356
QY 5161 agtcacactaagaagaatcaatgactaaatgaaaaaactaccaaagtatctctacagagcttca 5220
Db 5357 agtcacactaagaagaatcaatgactaaatgaaaaaactaccaaagtatctctacagagcttca 5416
QY 5221 cagccagacctgaotggacaccccacaaagctccaagccacagcccaagaagaagtcctcagg 5280
Db 5417 cagccagaccttagtgacaccccacaaagctccaagccacagcccaagaagaagtcctcagg 5476
QY 5281 aagcgacagactgaagaagaatttttagcattttaggaacaaacgcgcctcagcaggcaaa 5340
Db 5477 aagcgacagactgaagaagaatttttagcattttaggaacaaacgcgcctcagcaggcaaa 5536
QY 5341 gccatgcacacacccccaaaccagcagtaggtgaaagagaagaacatcaacgctttttggga 5400
Db 5537 gccatgcacacacccccaaaccagcagtaggtgaaagagaagaacatcaacgctttttggga 5596
QY 5401 actccagtgacagaactggacagccaggaattttacctggcagcaatagacggctacaa 5460
Db 5597 actccagtgacagaactggacagccaggaattttacctggcagcaatagacggctacaa 5556
QY 5461 actcgtlaaggaaaagcccaggtctctagaagaactgactggcttcagagagagcttttccag 5520
Db 5657 actcgtlaaggaaaagcccaggtctctagaagaactgactggcttcagagagagcttttccag 5716
QY 5521 acacctgcaactgataacccccacagctgatagaaaaactaccaaaaaaactctctgcaca 5580
Db 5717 acacctgcaactgataacccccacagctgatagaaaaactaccaaaaaaactctctgcaca 5776
QY 5581 tctccgcaatcagacccagcggaacccccacacaaacaaagcaacggcccaagagaagc 5640
Db 5777 tctccgcaatcagacccagcggaacccccacacaaacaaagcaacggcccaagagaagc 5836
QY 5641 ctcaaaaaagcagactagaggagaatttttagcatttcagggaacttaacacacatcagca 5700
Db 5837 ctcaaaaaagcagactagaggagaatttttagcatttcagggaacttaacacacatcagca 5896
QY 5701 ggcaagcccatgcacacgcctaaagcagcagtaggtgaaagagaagaacatcaacacattt 5760
Db 5897 ggcaagcccatgcacacgcctaaagcagcagtaggtgaaagagaagaacatcaacacattt 5956
QY 5761 gtgggactccagtggagaaaactggcctgctaggaatttacctggcagcagcagagacgg 5820
Db 5957 gtgggactccagtggagaaaactggcctgctaggaatttacctggcagcagcagagacgg 6016
QY 5821 ccacaaactcctaaagaaaaggcccaagggtcttagaagatctggtggttcaaaagagctc 5880
Db 6017 ccacaaactcctaaagaaaaggcccaagggtcttagaagatctggtggttcaaaagagctc 6076
QY 5881 ttccagacccaggtcacactgagggaatcaatgacccgatgacaaaatcacagaaagtatcc 5940
Db 6077 ttccagacccaggtcacactgagggaatcaatgacccgatgacaaaatcacagaaagtatcc 6136
QY 5941 tgcaaatctccacaacacagacccagtcacaaaccccaacagctcccaagcaacgactcaag 6000
Db 6137 tgcaaatctccacaacacagacccagtcacaaaccccaacagctcccaagcaacgactcaag 6196
QY 6001 atatacttgggaaaagttaggtgtgaaagaagaggttctccagctcggaagctcacacag 6060
Db 6197 atatacttgggaaaagttaggtgtgaaagaagaggttctccagctcggaagctcacacag 6256
QY 6061 acgtcagggaagaccacacagacacacagagagacagcagggagatgggaagagacacaaa 6120
Db 6257 acgtcagggaagaccacacagacacacagagagacagcagggagatgggaagagacacaaa 6316
QY 6121 ggcgttlaaggaatctcaaaagcagatgctggaccacgacaaactatggaactggagtgagag 6180
Db 6317 ggcgttlaaggaatctcaaaagcagatgctggaccacgacaaactatggaactggagtgagag 6376

QY 6181 aggtggccaaagacaccttaagaaagagcccaatactactagaaagacctggccggcttcaaa 6240
Db 6377 aggtggccaaagacaccttaagaaagagcccaatactactagaaagacctggccggcttcaaa 6436
QY 6241 gagctctccagacacacagaccactgaggaatcaacaactgatgacaaaactaccaaa 6300
Db 6437 gagctctccagacacacagaccactgaggaatcaacaactgatgacaaaactaccaaa 6496
QY 6301 atagcctgcaaatctccaccaccagaaatcaatggacactccaacaagcacaaggaggcg 6360
Db 6497 atagcctgcaaatctccaccaccagaaatcaatggacactccaacaagcacaaggaggcg 6556
QY 6361 cccaaaaacacctttgggaaaaagggatagtgaaagagctctcagccctgaagcagctc 6420
Db 6557 cccaaaaacacctttgggaaaaagggatagtgaaagagctctcagccctgaagcagctc 6616
QY 6421 acacagaccacacacagacaaagtaccagagagatgagataaaggcatcaacgctgttc 6480
Db 6617 acacagaccacacacagacaaagtaccagagagatgagataaaggcatcaacgctgttc 6676
QY 6481 agggaaactgcaaaacagaaactggaccagcagcaagtgtaactggtagcaagaggcag 6540
Db 6677 agggaaactgcaaaacagaaactggaccagcagcaagtgtaactggtagcaagaggcag 6736
QY 6541 ccaagaaactctaaagggaaaaagcccaaccccttagaagacttggtggttgaagagctc 6600
Db 6737 ccaagaaactctaaagggaaaaagcccaaccccttagaagacttggtggttgaagagctc 6796
QY 6601 ttccagacaccagtatgactgacaagcccacgactcacgagaaaactaccaaaatagcc 6660
Db 6797 ttccagacaccagtatgactgacaagcccacgactcacgagaaaactaccaaaatagcc 6856
QY 6661 tgcagatctccacacacagaccagtggtgtaccccaacaatcttcaagccacagtcctcaag 6720
Db 6857 tgcagatctccacacacagaccagtggtgtaccccaacaatcttcaagccacagtcctcaag 6916
QY 6721 agaaactctcaggaagacgacgtagagagaagaatctcttagcaactcagaaacagacaoca 6780
Db 6917 agaaactctcaggaagacgacgtagagagaagaatctcttagcaactcagaaacagacaoca 6976
QY 6781 tcaataggaagactatgacacaccccaacccagcagagatgatgagaagacatgaa 6840
Db 6977 tcaataggaagactatgacacaccccaacccagcagagatgatgagaagacatgaa 7036
QY 6841 gcatttatggaaactccagtcagaaattggacctgcaggaaatttacctggcagcaaa 6900
Db 7037 gcatttatggaaactccagtcagaaattggacctgcaggaaatttacctggcagcaaa 7096
QY 6901 agatggccacaactctctaaggaagggcccgagctcttagaagacctggctggttcaaa 6960
Db 7097 agatggccacaactctctaaggaagggcccgagctcttagaagacctggctggttcaaa 7156
QY 6961 gagctctccagacacacagcactgacaagcccacgactgatgagaaactaccaaaaata 7020
Db 7157 gagctctccagacacacagcactgacaagcccacgactgatgagaaactaccaaaaata 7216
QY 7021 gcttgcaaatctccacaaccagaccagtggaaccccagcaagcacaagcaacggccc 7080
Db 7217 gcttgcaaatctccacaaccagaccagtggaaccccagcaagcacaagcaacggccc 7276
QY 7081 aagagaacctcaggaagcagacgtagaggaagaatttttagcactcaggaacgaaca 7140
Db 7277 aagagaacctcaggaagcagacgtagaggaagaatttttagcactcaggaagcaaca 7336
QY 7141 coactcagcagcaagccatggacaccccaacccagcagtaagtgtagaaaaaatatc 7200
Db 7337 coactcagcagcaagccatggacaccccaacccagcagtaagtgtagaaaaaatatc 7396
QY 7201 aacacatttggaaactccagtgacagaaactggacctggtaggaatttacctggcagc 7260
Db 7397 aacacatttggaaactccagtgacagaaactggacctggtaggaatttacctggcagc 7456
QY 7261 aagagacagccacagactctctaaggaagggctgagggctctagaggacctggttggcttc 7320

Db 7457 aagagacagccacagactccttaagaaaagctgaggctctagaggacctggttggcttc 7516
QY 7321 aaagaactctccagacacccaggtcacactgagggaatcaatgactgatgacaaaatcaca 7380
Db 7517 aaagaactctccagacacccaggtcacactgagggaatcaatgactgatgacaaaatcaca 7576
QY 7381 gaagtactctgataatctccacagccagtgatctcaaaacctcaagaagctcccaagcaa 7440
Db 7577 gaagtactctgataatctccacagccagtgatctcaaaacctcaagaagctcccaagcaa 7636
QY 7441 aggtctcaagataaccctggtgaaagtggacatgaaagaagagccctcagcagtcagcaag 7500
Db 7637 aggtctcaagataaccctggtgaaagtggacatgaaagaagagccctcagcagtcagcaag 7696
QY 7501 ctcacacggacatcaggggagactacgcaaacacacacagagcccaacaggagatagtaag 7560
Db 7697 ctcacacggacatcaggggagactacgcaaacacacacagagcccaacaggagatagtaag 7756
QY 7561 agcatcaaaagcgtttaaggagtctcctcaaaagcagatctctggaccagcagcaagttaact 7620
Db 7757 agcatcaaaagcgtttaaggagtctcctcaaaagcagatctctggaccagcagcaagttaact 7816
QY 7621 ggtagcagaggcagctgagaactcgttaaggaaaaagggcccgtgctctagaagacctggtt 7680
Db 7817 ggtagcagaggcagctgagaactcgttaaggaaaaagggcccgtgctctagaagacctggtt 7876
QY 7681 gacttcaaaagactctctcagcacagggtcacactgaagagtcactgaatgactattgacaaa 7740
Db 7877 gacttcaaaagactctctcagcacagggtcacactgaagagtcactgaatgactattgacaaa 7936
QY 7741 aacacaaaaattccctgcataatctcccaccagaaactaacagacactgcccacgagcaca 7800
Db 7937 aacacaaaaattccctgcataatctcccaccagaaactaacagacactgcccacgagcaca 7996
QY 7801 aagaatgccccaaagacacgttcccggaaaaaggttaaaagagaggtctctcagcagttgag 7860
Db 7997 aagaatgccccaaagacacgttcccggaaaaaggttaaaagagaggtctctcagcagttgag 8056
QY 7861 aggtctcacgcaaaactcagggcaagcacaacacacacacaaagaacccagcagcaggtgat 7920
Db 8057 aggtctcacgcaaaactcagggcaagcacaacacacacacaaagaacccagcagcaggtgat 8116
QY 7921 gagggcatcaaatgattgaagcaacgtgcaaaagagaacacccacccagtagaagaggaa 7980
Db 8117 gagggcatcaaatgattgaagcaacgtgcaaaagagaacacccacccagtagaagaggaa 8176
QY 7981 cccagcaggaagcccaagcagcaactaaggaaaaagggcccaacccctggaaagacctggcc 8040
Db 8177 cccagcaggaagcccaagcagcaactaaggaaaaagggcccaacccctggaaagacctggcc 8236
QY 8041 ggcttcacagagctctctgaaaactcaggtcacactcagggaatcactgactgctggcaaa 8100
Db 8237 ggcttcacagagctctctgaaaactcaggtcacactcagggaatcactgactgctggcaaa 8296
QY 8101 gccactaaaaatccctgcgaaatctcccacactagaagtggttagacacacacagcagcaca 8160
Db 8297 gccactaaaaatccctgcgaaatctcccacactagaagtggttagacacacacagcagcaca 8356
QY 8161 aagaggcatctcagagacacgtgtgcagaaggtacaagtaaaagaagagacctcagcagctc 8220
Db 8357 aagaggcatctcagagacacgtgtgcagaaggtacaagtaaaagaagagacctcagcagctc 8416
QY 8221 aagttcacacaaaactcaggggaaaaccaggtgcagacaaaagaacccagcaggtgaagat 8280
Db 8417 aagttcacacaaaactcaggggaaaaccaggtgcagacaaaagaacccagcaggtgaagat 8476
QY 8281 aaaggcatcaaaagcattgaaggaatctgcacacacagacacccgctccagcagcagtgta 8340
Db 8477 aaaggcatcaaaagcattgaaggaatctgcacacacagacacccgctccagcagcagtgta 8536
QY 8341 actggcagcagagacgcccacagagccacccagggaaaagtgcaccaagccatagaagaccta 8400

Db 8537 actggcagcaggagagcgccaaagcaccacccagggaagtgtcccaagccatagagaccta 8596
QY 8401 gctggcttcaaaagacccagcagcaggttcacactgaagaatacaatgactgatgacaaaacc 8460
Db 8597 gctggcttcaaaagacccagcagcaggttcacactgaagaatacaatgactgatgacaaaacc 8556
QY 8461 actaaataacctgcaaatcatcaccagnactagaagacacccgcaacaagctcaaaagaga 8520
Db 8657 actaaataacctgcaaatcatcaccagactagaagacacccgcaacaagctcaaaagaga 8716
QY 8521 cggccagacacgtgcccagaagaagttagaagtgaaggagagctgttagcagttggcaag 8580
Db 8717 cggccagacacgtgcccagaagaagttagaagtgaaggagagctgttagcagttggcaag 8776
QY 8581 ctacacaaacctcaggggagacacgcacacgcgacaagagcggtaggtgagggcaaa 8640
Db 8777 ctacacaaacctcaggggagacacgcacacgcgacaagagcggtaggtgagggcaaa 8836
QY 8641 ggcacgaagcatttaagcaacctgcaaaagcggaacgtggacgcagaagaatgtaattggc 8700
Db 8837 ggcacgaagcatttaagcaacctgcaaaagcggaacgtggacgcagaagaatgtaattggc 8896
QY 8701 agcagagacagccaagacacctaaggaagaaagggcccaacctgggaagacctggccagc 8760
Db 8897 agcagagacagccaagacacctaaggaagaaagggcccaacctgggaagacctggccagc 8956
QY 8761 ttccaagagctctcaaacaccagcgccacactgaggaactggcgaatgtgtgctgctgat 8820
Db 8957 ttccaagagctctcaaacaccagcgccacactgaggaactggcgaatgtgtgctgctgat 9016
QY 8821 agctttacaagcgtccaaagcaaacacactgcagctggaaaacctctaaaaatatccaga 8880
Db 9017 agctttacaagcgtccaaagcaaacacactgcagctggaaaacctctaaaaatatccaga 9076
QY 8881 agagttcttcggcccttaaaagtagaacccgtggagacgtggttaagcaccagagaccct 8940
Db 9077 agagttcttcggcccttaaaagtagaacccgtggagacgtggttaagcaccagagaccct 9136
QY 8941 gtaaatccaaagcaaaagcaaaccttcctgcgcccccactgccttcagaggggaggt 9000
Db 9137 gtaaatccaaagcaaaagcaaaccttcctgcgcccccactgccttcagaggggaggt 9196
QY 9001 ggcacaagatggaagcgtcaacgggaaccaaagaggtcgctgcatgcagcaccagaggaa 9060
Db 9197 ggcacaagatggaagcgtcaacgggaaccaaagaggtcgctgcatgcagcaccagaggaa 9256
QY 9061 attgtgagagagctgcagccagcaagaagcagagaggttgcctccagggcaagagcaaa 9120
Db 9257 attgtgagagagctgcagccagcaagaagcagagaggttgcctccagggcaagagcaaa 9316
QY 9121 tcattccgaacccgtggtcatcatgaagagaagtttgagagacttctgcacaaagaattgaa 9180
Db 9317 tcattccgaacccgtggtcatcatgaagagaagtttgagagacttctgcacaaagaattgaa 9376
QY 9181 cctgcggaagagctgaacgcgaacgcacatgaaacccaacaaagaggaacacaaattacaa 9240
Db 9377 cctgcggaagagctgaacgcgaacgcacatgaaacccaacaaagaggaacacaaattacaa 9436
QY 9241 gactcgtctccctgaaataaagggaataatccctgcgtccagcagccaagataagactgag 9300
Db 9437 gactcgtctccctgaaataaagggaataatccctgcgtccagcagccaagataagactgag 9496
QY 9301 gcagaacagcaataaactgaggtctttgtattagcagaagaataagaataaacagaaat 9360
Db 9497 gcagaacagcaataaactgaggtctttgtattagcagaagaataagaataaacagaaat 9556
QY 9361 gaaaagaagcccatgaagacacctcccagagatggacattcagaatccagatgatggagcc 9420
Db 9557 gaaaagaagcccatgaagacacctcccagagatggacattcagaatccagatgatggagcc 9616
QY 9421 cggaaacccatccttagagacaaaagtccactgagacaaaaggtcttaggtctgactaga 9480
Db 9617 cggaaacccatccttagagacaaaagtccactgagacaaaaggtcttaggtctgactaga 9676

QY 9481 cagaatgagagctcccagcctaaagtgtggcagagagagcgagggcagagaagtgcgaag 9540
Db 9677 cagaatgagagctcccagcctaaagtgtggcagagagagcgagggcagagaagtgcgaag 9736
QY 9541 gttctcatgcagaatcagaaaagggaaagagcagggaattccagactccatgtgcctg 9600
Db 9737 gttctcatgcagaatcagaaaagggaaagagcagggaattccagactccatgtgcctg 9796
QY 9601 agatcaagaagaagaaaaagccagcctgcagcaagcacitttgagagcaaatctgtgcag 9660
Db 9797 agatcaagaagaagaaaaagccagcctgcagcaagcacitttgagagcaaatctgtgcag 9856
QY 9661 agagtcaacgcgagtgctcaagaggtgtgcagaaaaatccaaaagagctgaggacaatgtg 9720
Db 9857 agagtcaacgcgagtgctcaagaggtgtgcagaaaaatccaaaagagctgaggacaatgtg 9916
QY 9721 tgtgtcaagaagaataacaaccagaaagtcataagcagtgaaagata 9766
Db 9917 tgtgtcaagaagaataacaaccagaaagtcataagcagtgaaagata 9962

RESULT 2
AAA64658
ID AAA64658 standard; cDNA; 12515 BP.
XX
AC AAA64658;
XX
DT 02-JAN-2001 (first entry)
XX
DE cDNA sequence encoding the antigen of monoclonal antibody Ki-67.
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW antibody Ki-67; Ss.
OS Homo sapiens.
XX
XX WO2000050595-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04929.
XX
PR 25-FEB-1999; 99US-0257417.
XX
PA (GOUT/) GOUT I.
PA (RODN/) RODNIN N.
PA (FILO/) FILONENKO V.
PA (MATS/) MATSUKA G.
PA (SCAN/) SCANLAN M.
PA (OLDL/) OLD L.
PA (BILY/) BILYNSKY B.
XX
PI Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
PI Bilynsky B;
XX
DR WPI; 2000-572092/53.
XX
PT Novel isolated nucleic acid molecules for diagnosing and treating -
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -
XX
PS Claim 17; Page 77-81; 94pp; English.
XX
CC The specification describes polynucleotides which are associated with
CC melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC diagnosing and treating a patient with melanoma, thyroid tumour,
CC rectal cancer, lung cancer, breast cancer or colon cancer. The
CC present sequence represents a polynucleotide of the invention.
XX
SQ Sequence 12515 BP; 4164 A; 3050 C; 2928 G; 2373 T; 0 other;

Query Match 94.5%; Score 9232; DB 21; Length 12515;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 9762; Conservative 0; Mismatches 1; Indels 6; Gaps 2;

Qy 1 atgtggccacgagacgctggttactatcaaaaggagcggtgcagcggtcccccattt 60
Db 197 atgtggccacgagacgctggttactatcaaaaggagcggtgcagcggtcccccattt 256
Qy 61 cccctgagccctcagcacctgctgtttggaagggtattgaatgagcatcogtatccag 120
Db 257 cccctgagccctcagcacctgctgtttggaagggtattgaatgagcatcogtatccag 316
Qy 121 ctctctgttgttgcataaacacattgcgaagtgaattcccatgagcagaggaatatta 180
Db 317 ctctctgttgttgcataaacacattgcgaagtgaattcccatgagcagaggaatatta 376
Qy 181 cataatttcagttccataaacatccacacagaataatggaatggtgttatgagagcctgta 240
Db 377 cataatttcagttccataaacatccacacagaataatggaatggtgttatgagagcctgta 436
Qy 241 cggctaaaacatggagatgtaataactattattgatctgttcttcaggtatgaaatgaa 300
Db 437 cggctaaaacatggagatgtaataactattattgatctgttcttcaggtatgaaatgaa 496
Qy 301 agtcttcagaatggaaaggaaagtcgaactgaatttccaaagaaaaatacgtgaacagagcca 360
Db 497 agtcttcagaatggaaaggaaagtcgaactgaatttccaaagaaaaatacgtgaacagagcca 556
Qy 361 gcaactcgtgtctcaagatctagcttctctctgtaccctgatgagaagcgtcaagattcc 420
Db 557 gcaactcgtgtctcaagatctagcttctctctgtaccctgatgagaagcgtcaagattcc 616
Qy 421 aaggcctattcaaaatcacgaaagaaagtcttcagaaatcctcaggtacatatacag 480
Db 617 aaggcctattcaaaatcacgaaagaaagtcttcagaaatcctcaggtacatatacag 676
Qy 481 aatgtcaagaagacagtlaccgcagatgactcaaaagacagtggttgctcagggaaacaact 540
Db 677 aatgtcaagaagacagtlaccgcagatgactcaaaagacagtggttgctcagggaaacaact 736
Qy 541 aatgttcatctcagaacatgctggaatgtaatggcagaataatgcagctgatccattct 600
Db 737 aatgttcatctcagaacatgctggaatgtaatggcagaataatgcagctgatccattct 796
Qy 601 ggggattttaagaaatttcagcgttaaatagtgagccgttatggagaattgagct 660
Db 797 ggggattttaagaaatttcagcgttaaatagtgagccgttatggagaattgagct 856
Qy 661 gtccactacacaatgtcttgacaatagcaaaaaaatgaatctcccttttggagctt 720
Db 857 gtccactacacaatgtcttgacaatagcaaaaaaatgaatctcccttttggagctt 916
Qy 721 tatgagtcagtgaaagaagagtgatgtataaatcacaaaaagaaaaatgtctcagat 780
Db 917 tatgagtcagtgaaagaagagtgatgtataaatcacaaaaagaaaaatgtctcagat 976
Qy 781 tgtagaaatctggattacaactgattacgacaagagaagaagtgtgatgttita 840
Db 977 tgtagaaatctggattacaactgattacgacaagagaagaagtgtgatgttita 1036
Qy 841 caggggagacccaactgttggtctcgcgtgaagtcaagaccaaaaatctgggtggagcggc 900
Db 1037 caggggagacccaactgttggtctcgcgtgaagtcaagaccaaaaatctgggtggagcggc 1096
Qy 901 cagcgtgtggcagagcctgttccactgaacaaagagcttgaccagaaacaaaggggagggga 960
Db 1097 cagcgtgtggcagagcctgttccactgaacaaagagcttgaccagaaacaaaggggagggga 1156
Qy 961 agagcgtggagctgtgttcagactcccaagcagctgtgggcccagctttctctctat 1020
Db 1157 agagcgtggagctgtgttcagactcccaagcagctgtgggcccagctttctctctat 1216

Qy 1021 gagcgggtaaaaatgaagaccctgtacaattatcacagcaacaaaattctcccaaaaa 1080
Db 1217 gagcgggtaaaaatgaagaccctgtacaattatcacagcaacaaaattctcccaaaaa 1276
Qy 1081 cataagaacaaagacctgtataactactggtagaagaatctgtgaatctgggtaaaaagt 1140
Db 1277 cataagaacaaagacctgtataactactggtagaagaatctgtgaatctgggtaaaaagt 1336
Qy 1141 gaaggcttcaaggctgtgtgataaaaactctactccaggaagcttttcaactagaaatcga 1200
Db 1337 gaaggcttcaaggctgtgtgataaaaactctactccaggaagcttttcaactagaaatcga 1396
Qy 1201 acaccagctaaagtgtgaagatgcagctgactctgcactaaagccagaaaaatctctctcc 1260
Db 1397 acaccagctaaagtgtgaagatgcagctgactctgcactaaagccagaaaaatctctctcc 1456
Qy 1261 aaaaacagaggaagtattctctacagatgtggaagtctctcctacggaaaaactgaaattcac 1320
Db 1457 aaaaacagaggaagtattctctacagatgtggaagtctctcctacggaaaaactgaaattcac 1516
Qy 1321 aatgagccatttttaactctgtggtcactcaagtgtgagaggaatcccaaaagattccc 1380
Db 1517 aatgagccatttttaactctgtggtcactcaagtgtgagaggaatcccaaaagattccc 1576
Qy 1381 ctccagcaagcctgagaatttgggcactacagctggacagatgtctctgggttacctggt 1440
Db 1577 ctccagcaagcctgagaatttgggcactacagctggacagatgtctctgggttacctggt 1636
Qy 1441 cttaagtcaagttgatatacaaaactttggtgattccattaatgagagtgaggaatacct 1500
Db 1637 cttaagtcaagttgatatacaaaactttggtgattccattaatgagagtgaggaatacct 1696
Qy 1501 ttgaaaaagaagcgtgtgtcctttggtggccactaaagccctgaactatttggatgaaaaa 1560
Db 1697 ttgaaaaagaagcgtgtgtcctttggtggccactaaagccctgaactatttggatgaaaaa 1756
Qy 1561 ttgctctctaatcacgctctcaaaaggggagagcccccacaaagaagaagtctctggta 1620
Db 1757 ttgctctctaatcacgctctcaaaaggggagagcccccacaaagaagaagtctctggta 1816
Qy 1621 atgcacactccacctgtcctgaagaaaaatcatcaaggaaacagcctcaaacatcaggaaaa 1680
Db 1817 atgcacactccacctgtcctgaagaaaaatcatcaaggaaacagcctcaaacatcaggaaaa 1876
Qy 1681 caagagtcaggttccagaaatccatgtggaagtgaaggccaaagcttggttataagccct 1740
Db 1877 caagagtcaggttccagaaatccatgtggaagtgaaggccaaagcttggttataagccct 1936
Qy 1741 ccagctccttagtctcaggaaaaactccagttgccagtgatcaacgcccgtaggtcctgcaaa 1800
Db 1937 ccagctccttagtctcaggaaaaactccagttgccagtgatcaacgcccgtaggtcctgcaaa 1996
Qy 1801 acagccctgtctcagcagcaaaaatctcagacagaggttccctaaagagagagagaaaaa 1860
Db 1997 acagccctgtctcagcagcaaaaatctcagacagaggttccctaaagagagagagaaaaa 2056
Qy 1861 gtggcaacctgcttcaaaaagagtgctctatcagccgaagtcaacatgatatatttccag 1920
Db 2057 gtggcaacctgcttcaaaaagagtgctctatcagccgaagtcaacatgatatatttccag 2116
Qy 1921 atgatgtgtccaaaagagaagtgtggtctcgggaagcaaatctgattgttgcaaaatca 1980
Db 2117 atgatgtgtccaaaagagaagtgtggtctcgggaagcaaatctgattgttgcaaaatca 2176
Qy 1981 tgggcagatgtagtataaaacttgggtcaaaaacaaacaaactaaagtataaaataggt 2040
Db 2177 tgggcagatgtagtataaaacttgggtcaaaaacaaacaaactaaagtataaaataggt 2236
Qy 2041 cctcaaaaggtcaatgaacaaaagcgaagagacctgtactccaaagagcctgtgggc 2100
Db 2237 cctcaaaaggtcaatgaacaaaagcgaagagacctgtactccaaagagcctgtgggc 2296
Qy 2101 gaagttcacagtcatttagtacagggccacgcaaacctctcttctaccataataataggg 2160

Db 2297 gaagttcacagtcgaatttagtacaggccacgcgaactctccttctgaccataataataggg 2356
QY 2161 aaagtcatactgaaaaagtacatgtgcctgtcgcacctacagagtgtctcaacaacttc 2220
Db 2357 aaagtcatactgaaaaagtacatgtgcctgtcgcacctacagagtgtctcaacaacttc 2416
QY 2221 atttccaaacaaaaattggaactttaaggaaagtcttctcaggaatagctgaaatgtttcaag 2280
Db 2417 atttccaaacaaaaattggaactttaaggaaagtcttctcaggaatagctgaaatgtttcaag 2476
QY 2281 acccagtgaaagagcaaccgcagttgcagaagcacatgtccatctctatttccaaattcca 2340
Db 2477 acccagtgaaagagcaaccgcagttgcagaagcacatgtccatctctatttccaaattcca 2536
QY 2341 gagaatttgccttgaaaaacagttctcaggaactgattcaggagaagaactctctgctcccc 2400
Db 2537 gagaatttgccttgaaaaacagttctcaggaactgattcaggagaagaactctctgctcccc 2596
QY 2401 acctcagagagtttggaggaaaatgtgttcttcagtgccacagaaatgcagcaaaacagcca 2460
Db 2597 acctcagagagtttggaggaaaatgtgttcttcagtgccacagaaatgcagcaaaacagcca 2656
QY 2461 tctgataaatgtctgcagaccctcccttaagacggcagtgatttagagaaaaatgaaac 2520
Db 2657 tctgataaatgtctgcagaccctcccttaagacggcagtgatttagagaaaaatgaaac 2716
QY 2521 gtagcaaaaacgcccaggaaacactcaaaaatgacttctctggagacaaaactcttcagat 2580
Db 2717 gtagcaaaaacgcccaggaaacactcaaaaatgacttctctggagacaaaactcttcagat 2776
QY 2581 actgagacagagccttcaaaaacagttatccactgttaaacaggtcaggaaaggtctacagag 2640
Db 2777 actgagacagagccttcaaaaacagttatccactgttaaacaggtcaggaaaggtctacagag 2836
QY 2641 ttacagaaatacacagaagtacactgtggaagtaagagtgaaagaaacaaatacacagaatt 2700
Db 2837 ttacagaaatacacagaagtcacactgtggaagtaagagtgaaagaaacaaatacacagaatt 2896
QY 2701 gttgagtgcatctctaaaaagaggtcagagggcaacactactacaaacaaaggagagaagga 2760
Db 2897 gttgagtgcatctctaaaaagaggtcagagggcaacactactacaaacaaaggagagaagga 2956
QY 2761 gagatgaagaaatagagaagacccttttgagacataaaaggaaaaatttgaattaaagaa 2820
Db 2957 gagatgaagaaatagagaagacccttttgagacataaaaggaaaaatttgaattaaagaa 3016
QY 2821 aacgatgaagaagatgaagcaatgaagagatcaagaaacttggggcgagaatgtgcacca 2880
Db 3017 aacgatgaagaagatgaagcaatgaagagatcaagaaacttggggcgagaatgtgcacca 3076
QY 2881 atgtctgacctgacagacctcaagagcttgacctgtatcacagaactctatgaagaacagcca 2940
Db 3077 atgtctgacctgacagacctcaagagcttgacctgtatcacagaactctatgaagaacagcca 3136
QY 2941 cgtggcagaatctctccaaaccccaagaatcatgccaaggcaccaaagagtgagaagggc 3000
Db 3137 cgtggcagaatctctccaaaccccaagaatcatgccaaggcaccaaagagtgagaagggc 3196
QY 3001 aaaaactactaaaatgccctgccagttcattacaacgaagaacaaataaaccacacacac 3060
Db 3197 aaaaactactaaaatgccctgccagttcattacaacccgaacaaataaaccacacacacac 3256
QY 3061 acaaaacaacagttgaaggcatccctgggggaaagtaggtgtgaaagaagagctcccttagca 3120
Db 3257 acaaaacaacagttgaaggcatccctgggggaaagtaggtgtgaaagaagagctcccttagca 3316
QY 3121 gtcggcaagtccacgcagctcaggggagaccacgcacacgcacagagagccagcagga 3180
Db 3317 gtcggcaagtccacgcagctcaggggagaccacgcacgcacagagagccagcagga 3376
QY 3181 gatggcaagagcatcagaagctttaaggaggtctccaaagcagatcttgaccacagagcc 3240
Db 3377 gatggcaagagcatcagaagctttaaggaggtctccaaagcagatcttgaccacagagcc 3436

QY 3241 cgtgtaactggaatgaagaagtggccaaagaaacgcctaaagaaagagggcccgactactagaa 3300
Db 3437 cgtgtaactggaatgaagaagtggccaaagaaacgcctaaagaaagagggcccgactactagaa 3496
QY 3301 gacctggctggcttcaaaagagctctccagacacccgggtccctctcaggaatcaatgact 3360
Db 3497 gacctggctggcttcaaaagagctctccagacacccgggtccctctcaggaatcaatgact 3556
QY 3361 gatgagaaaactaccaaaatagcctgcaaatct---ccaccacgaatacagtgggacact 3417
Db 3557 gatgagaaaactaccaaaatagcctgcaaatctcccccaccacgaatacagtgggacact 3616
QY 3418 caacaagcacaaagcaatggcctaaagaaagtctcaggaagacagatgtagaggaagaa 3477
Db 3617 caacaagcac---gcaatggcctaaagaaagtctcaggaagacagatgtagaggaagaa 3673
QY 3478 ttcttagcactcaggaaaactaacaccatcagcagggaagccaatgtcttacgcgccaaacca 3537
Db 3674 ttcttagcactcaggaaaactaacaccatcagcagggaagccaatgtcttacgcgccaaacca 3733
QY 3538 gcagagagtgatgagaaagacattaaagcatttatgggaactccagtgcaagaactggac 3597
Db 3734 gcagagagtgatgagaaagacattaaagcatttatgggaactccagtgcaagaactggac 3793
QY 3598 ctggcgaggaactttacctgtgcagcaaaagacagctacagactcctaaaggaaaagggcccg 3657
Db 3794 ctggcgaggaactttacctgtgcagcaaaagacagctacagactcctaaaggaaaagggcccg 3853
QY 3658 gctctagaagacctggctggctttaagagctcttccagactcctgggtcacaccgggaa 3717
Db 3854 gctctagaagacctggctggctttaagagctcttccagactcctgggtcacaccgggaa 3913
QY 3718 ttagtgtcgtcgtgtaaaaccactaaaaaccctgcagctctccacagtcagaccagtg 3777
Db 3914 ttagtgtcgtcgtgtaaaaccactaaaaaccctgcagctctccacagtcagaccagtg 3873
QY 3778 gacaccccaacagcaacaaagcaacgaccccaagagaglatcaggaagcagatgtagag 3837
Db 3974 gacaccccaacagcaacaaagcaacgaccccaagagaglatcaggaagcagatgtagag 4033
QY 3838 gagaaactcttagctgctgagggaaatctaattgcactcagcagggaacccaatgcacagcct 3897
Db 4034 gagaaactcttagctgctgagggaaatctaattgcactcagcagggaacccaatgcacagcct 4093
QY 3898 aaacctcagtaggtgaagaaaaagacatcatatttctggtggaactccagtgccagaa 3957
Db 4094 aaacctcagtaggtgaagaaaaagacatcatatttctggtggaactccagtgccagaa 4153
QY 3958 ctggacctgacagagaaacttaaccggcagcagagagcgcacaaactccttaagggaagag 4017
Db 4154 ctggacctgacagagaaacttaaccggcagcagagagcgcacaaactccttaagggaagag 4213
QY 4018 gccaggtctctggaagacctgactggcttttaaagagctcttccagaccctctgctact 4077
Db 4214 gccaggtctctggaagacctgactggcttttaaagagctcttccagaccctctgctact 4273
QY 4078 gaagaagcagtggtgtctggcaaaactactaaaatgccctgcgaatcttctccaccagaa 4137
Db 4274 gaagaagcagtggtgtctggcaaaactactaaaatgccctgcgaatcttctccaccagaa 4333
QY 4138 tcagcgacacccccaacaaagcacaagaagcgaccccaagacaccttttgagaaaaagggac 4197
Db 4334 tcagcgacacccccaacaaagcacaagaagcgaccccaagacaccttttgagaaaaagggac 4393
QY 4198 gtacagaagagctctcagccctgaagagagctcacacagacatcaggggaaacacacac 4257
Db 4394 gtacagaagagctctcagccctgaagagagctcacacagacatcaggggaaacacacac 4453
QY 4258 acagaataaagtagccagaggtgaggaataaagacatcaacgcgttttagggaactgcacaa 4317
Db 4454 acagaataaagtagccagaggtgaggaataaagacatcaacgcgttttagggaactgcacaa 4513

QY 4318 cagaactggacccagcagcaagtgttaactggtagcaagaggcaccacaaactaaggaa 4377
DB 4514 cagaaactggacccagcagcaagtgttaactggttagcaagaggcaccacaaactaaggaa 4573
QY 4378 aaggcccaacccctagaagaactggctggctggaagaagctcttcacagaccagtatgc 4437
DB 4574 aaggcccaacccctagaagaactggctggctggaagaagctcttcacagaccagtatgc 4633
QY 4438 actgacaagcccaagactcacgagaagaaactaccaaaatagcctgcagatcacaaaccagac 4497
DB 4634 actgacaagcccaagactcacgagaagaaactaccaaaatagcctgcagatcacaaaccagac 4693
QY 4498 ccagtggcacacacaaagctccaagccacagctccaagagaagctctcaggaaaagtggac 4557
DB 4694 ccagtggcacacacaaagctccaagccacagctccaagagaagctctcaggaaaagtggac 4753
QY 4558 gtagaagaagaattcttcgcactcagggaagaaacacacacatcagcagggaagccatgcac 4617
DB 4754 gtagaagaagaattcttcgcactcagggaagaaacacacacatcagcagggaagccatgcac 4813
QY 4618 acacccaaacagcagtagtaagtgttgagaagaaactctacgcatttatggaaactccagt 4677
DB 4814 acacccaaacagcagtagtaagtgttgagaagaaactctacgcatttatggaaactccagt 4873
QY 4678 cagaaactggacctgacagagaacttaactggcagcaagagacggctacaaactccttaag 4737
DB 4874 cagaaactggacctgacagagaacttaactggcagcaagagacggctacaaactccttaag 4933
QY 4738 gaaaaggcccggtctctagaagacctggctggctttaaagagctcttcacagacagaggt 4797
DB 4934 gaaaaggcccggtctctagaagacctggctggctttaaagagctcttcacagacagaggt 4993
QY 4798 cacactgagggaatcaatgactaaacgataaaactccaaagtagcctgcaaatcttcacaa 4857
DB 4994 cacactgagggaatcaatgactaaacgataaaactccaaagtagcctgcaaatcttcacaa 5053
QY 4858 ccagacctagacaaaaacccagcaagctccaagcgacggctcaagacatccctctgggaaa 4917
DB 5054 ccagacctagacaaaaacccagcaagctccaagcgacggctcaagacatccctctgggaaa 5113
QY 4918 gtggcggtgaagaagaagctccttagcagttggcaagctccacagacatcaggagagact 4977
DB 5114 gtggcggtgaagaagaagctccttagcagttggcaagctccacagacatcaggagagact 5173
QY 4978 acacacacacacagagcccaacagggagatggttaagagatgaaagcatttatgaggtct 5037
DB 5174 acacacacacacagagcccaacagggagatggttaagagatgaaagcatttatgaggtct 5233
QY 5038 ccaagcgagatcttagactcagcagcaagcttaactgcagcaagaggcagctgagaact 5097
DB 5234 ccaagcgagatcttagactcagcagcaagcttaactgcagcaagaggcagctgagaact 5293
QY 5098 cctaagggaagctgtgaagtccttgaaagacgtggccggtctcatagagctcttcacagaca 5157
DB 5294 cctaagggaagctgtgaagtccttgaaagacgtggccggtctcatagagctcttcacagaca 5353
QY 5158 ccaagtcacactaagggaatcaatgactaataaaaaactaccaaagtatcctacagagct 5217
DB 5354 ccaagtcacactaagggaatcaatgactaataaaaaactaccaaagtatcctacagagct 5413
QY 5218 tcacagcgagacttagtggacaccccaacaaagctccaagggcacagcccaagagaagctc 5277
DB 5414 tcacagcgagacttagtggacaccccaacaaagctccaagggcacagcccaagagaagctc 5473
QY 5278 aggaagcagacactgaagagaatttttagcattttagaagaaacaaacccaatcagcaggg 5337
DB 5474 aggaagcagacactgaagagaatttttttagcattttagaagaaacaaacccaatcagcaggg 5533
QY 5338 aaagcctatgcacacaccccaacccagcagtaggtgaagagaagacatcaacacgtttttg 5397
DB 5534 aaagcctatgcacacaccccaacccagcagtaggtgaagagaagacatcaacacgtttttg 5593

QY 5398 ggaactccagtgcaagaactggacccagcaggaataattttacctggcagcaataagcggcta 5457
DB 5594 ggaactccagtgcaagaactggacccagcaggaataattttacctggcagcaataagcggcta 5653
QY 5458 caaactcgttaaggaaaaagggccaggctctagaagaactgactggcttcagagagcttttc 5517
DB 5654 caaactcgttaaggaaaaagggccaggctctagaagaactgactggcttcagagagcttttc 5713
QY 5518 cagacacactgcactgtataacccccacagctgatgaaaaactaccaaaaatactctgc 5577
DB 5714 cagacacactgcactgtataacccccacagctgatgaaaaactaccaaaaatactctgc 5773
QY 5578 aaatctcgcgaatcagaccccgagcagcccccacaaacacacaaagcaacggcccaagaga 5637
DB 5774 aaatctcgcgaatcagaccccgagcagcccccacaaacacacaaagcaacggcccaagaga 5833
QY 5638 agcctcaagaagaagcagcgttagaggagaatttttagcattcaggaaaaactaacaccatca 5697
DB 5834 agcctcaagaagaagcagcgttagaggagaatttttagcattcaggaaaaactaacaccatca 5893
QY 5698 gcaggcaaaagcccatgcacacgcttaaaagcagcagtaggtgaagagaagacatcaacaca 5757
DB 5894 gcaggcaaaagcccatgcacacgcttaaaagcagcagtaggtgaagagaagacatcaacaca 5953
QY 5758 ttgtggggactcccgtagaactggacctgctaggaaaattttacctggcagcagcagaga 5817
DB 5954 ttgtggggactcccgtagaactggacctgctaggaaaattttacctggcagcagcagaga 6013
QY 5818 cggccacaaactccttaaagaaaagcccaaggctctagaagaatctggctggcttcaaaagag 5877
DB 6014 cggccacaaactccttaaagaaaagcccaaggctctagaagaatctggctggcttcaaaagag 6073
QY 5878 ctctccagacacccaggttcacactgagaaatcaatgacccgtgcaaaaatcacagaatca 5937
DB 6074 ctctccagacacccaggttcacactgagaaatcaatgacccgtgcaaaaatcacagaatca 6133
QY 5938 tctgcaaatctcccaacacccagacccagtcacaaaccccaacaaagctccaagcaacgactc 5997
DB 6134 tctgcaaatctcccaacacccagacccagtcacaaaccccaacaaagctccaagcaacgactc 6193
QY 5998 aagatatccttggggaaaagtagtgtgaaaagagggtcctaccagtcggcgaagctcaca 6057
DB 6194 aagatatccttggggaaaagtagtgtgaaaagagggtcctaccagtcggcgaagctcaca 6253
QY 6058 cagacgtcagggaagaagcacacacagacagacagcagcagagatggaagaagcactc 6117
DB 6254 cagacgtcagggaagaagcacacacagacagacagcagcagagatggaagaagcactc 6313
QY 6118 aaagcgtttaagggaatctgcaaaagcagatgctggaccacagcaaaactatggaactgggagtg 6177
DB 6314 aaagcgtttaagggaatctgcaaaagcagatgctggaccacagcaaaactatggaactgggagtg 6373
QY 6178 gagaagtgcccaagaacacctaaggaagggcccaatcactagaagacactggccggcttc 6237
DB 6374 gagaagtgcccaagaacacctaaggaagggcccaatcactagaagacactggccggcttc 6433
QY 6238 aaagagctcttcacagacacccagacacacactgagggaatcaacaactgatgacaaaaactcc 6297
DB 6434 aaagagctcttcacagacacccagacacacactgagggaatcaacaactgatgacaaaaactcc 6493
QY 6298 aaaa tagcctgcgaatctccacccacagaatcaatggacactccaacagcacaagcagagg 6357
DB 6494 aaaa tagcctgcgaatctccacccacagaatcaatggacactccaacagcacaagcagagg 6553
QY 6358 cggcccaaaaacaccttttgggaaaaagggatagtggaagagctctcagccctgaagcag 6417
DB 6554 cggcccaaaaacaccttttgggaaaaagggatagtggaagagctctcagccctgaagcag 6613
QY 6418 ctcacagacacacacacacagacaaagtaccagagatgagataaaagggcatcaacgtg 6477
DB 6614 ctcacagacacacacacacagacaaagtaccagagatgagataaaagggcatcaacgtg 6573
QY 6478 ttcaagggaactgcaaaaacacagaaaactggaccacagcagcaagtgttaactggtagcagagg 6537

Db 6674 ttccagggaactgcataaacgaactggaccagcagcaagttaactggtagcaagagg 6733
QY cagccagaactccttaagggaagcccaaccctctagaagacttggctggttgaagag 6597
Db 6734 cagccagaactccttaagggaagcccaaccctctagaagacttggctggttgaagag 6793
QY ctcttcagacacccagtatgactgacaagcccacgactcacgagaaaactaccaaaata 6657
Db 6794 ctcttcagacacccagtatgactgacaagcccacgactcacgagaaaactaccaaaata 6853
QY gctcgagatctccaaaccagaccagtggttaccccaacaattcttcaaggccacagtc 6717
Db 6854 gctcgagatctccaaaccagaccagtggttaccccaacaattcttcaaggccacagtc 6913
QY aagagaagttctcaggaaaagcagcgttagaggaagaatccttagcactcaggaaaacgaaca 6777
Db 6914 aagagaagttctcaggaaaagcagcgttagaggaagaatccttagcactcaggaaaacgaaca 6973
QY ccatagtagggaagactatggacacaccccaaccagcaggaggtgatgagaagacatg 6837
Db 6974 ccatagtagggaagactatggacacaccccaaccagcaggaggtgatgagaagacatg 7033
QY aaagcatttatgggaactccagtcgagaaattggacctgcaggaataatttacctggcagc 6897
Db 7034 aaagcatttatgggaactccagtcgagaaattggacctgcaggaataatttacctggcagc 7093
QY aaagaatggccacaactccttaaggaaaagcccaggtcttagaagacctggctgcttc 6957
Db 7094 aaagaatggccacaactccttaaggaaaagcccaggtcttagaagacctggctgcttc 7153
QY aaagagctcttcagacacaccagcactgacaagcccacgactgatgagaaaactaccaa 7017
Db 7154 aaagagctcttcagacacaccagcactgacaagcccacgactgatgagaaaactaccaa 7213
QY atagctgcaaatctccaaaccagaccagtggaacaccccaagcagacaaagcaacgg 7077
Db 7214 atagctgcaaatctccaaaccagaccagtggaacaccccaagcagacaaagcaacgg 7273
QY cccaagagaactcctcaggaaaagcagcgttagaggaagaatttttagcactcaggaaaacga 7137
Db 7274 cccaagagaactcctcaggaaaagcagcgttagaggaagaatttttagcactcaggaaaacga 7333
QY acaccatcagcaggaaaagcca tggacaccccaaaaccagcagtaagttagagaaaaat 7197
Db 7334 acaccatcagcaggaaaagcca tggacaccccaaaaccagcagtaagttagagaaaaat 7393
QY atcaacacatttgtgaaactccagtcgagaaactggacctgctagaaaaatttacctggc 7257
Db 7394 atcaacacatttgtgaaactccagtcgagaaactggacctgctagaaaaatttacctggc 7453
QY agcaagagacagccacagactccttaaggaaaaggttgaggctctagagacctgggtggc 7317
Db 7454 agcaagagacagccacagactccttaaggaaaaggttgaggctctagagacctgggtggc 7513
QY ttcaagaactcttcagacacaccaggttcacactgaggaaatcaatgactgatgacaaaatc 7377
Db 7514 ttcaagaactcttcagacacaccaggttcacactgaggaaatcaatgactgatgacaaaatc 7573
QY acagaagatctcgttaaatctccacagccagagtcattcaaaaactcaagaagactccaag 7437
Db 7574 acagaagatctcgttaaatctccacagccagagtcattcaaaaactcaagaagactccaag 7633
QY caaagctcaagatacccttggtgaagtggacatgaagaagagcccttagcagtcagc 7497
Db 7634 caaagctcaagatacccttggtgaagtggacatgaagaagagcccttagcagtcagc 7693
QY aagctcacacggacatacagggggagactacgcaaacacacacagagccaacagagatag 7557
Db 7694 aagctcacacggacatacagggggagactacgcaaacacacacagagccaacagagatag 7753
QY aagagcatcaagcgttttaaggagttccaaagcagactcctggaccagcagcaagtga 7617
Db aagggcacgaaaagcattttaagcaacctgcaaaagcagcgtggaccagagaagtgttaatt 7893

Db 7754 aagagcatcaaaagcgtttaaggaggtctccaaagcagatccttggaaccagcagcaagtga 7813
QY actggtagcaggaggcagctgagaactcgtaaagaaaagcccgctctctagaagacctg 7677
Db 7814 actggtagcaggaggcagctgagaactcgttaagaaaaagcccgctctctagaagacctg 7873
QY gttgacttcaaaagagctcttctcagcaccagggtcacactgaagagtcacatgactattgac 7737
Db 7874 gttgacttcaaaagagctcttctcagcaccagggtcacactgaagagtcacatgactattgac 7933
QY aaaaaacacaaaaattccctgcataatctccccaccagaaataacagacactgccacagc 7797
Db 7934 aaaaaacacaaaaattccctgcataatctccccaccagaaataacagacactgccacagc 7993
QY acaaaagatgcccccaagacacgtccccaggaagaagtaaaagagggagctctcagcagtt 7857
Db 7994 acaaaagatgcccccaagacacgtccccaggaagaagtaaaagagggagctctcagcagtt 8053
QY gagagctcgcgcaaacatcagggcgaagcacacacacacacacacacacacacacacacac 7917
Db 8054 gagagctcgcgcaaacatcagggcgaagcacacacacacacacacacacacacacacac 8113
QY gatgaagggcatcaaaagtattgaagcaacgtgcataaagaagaaccccaaccagtagaagag 7977
Db 8114 gatgaagggcatcaaaagtattgaagcaacgtgcataaagaagaaccccaaccagtagaagag 8173
QY gaaccagcaggaagggcccaagcacccttaaggaaaagggcccaaccctctgggaagacctg 8037
Db 8174 gaaccagcaggaagggcccaagcacccttaaggaaaagggcccaaccctctgggaagacctg 8233
QY gccgggttcacagagctctctgaacatcaggttcacactcaggaatcactgactgctggc 8097
Db 8234 gccgggttcacagagctctctgaacatcaggttcacactcaggaatcactgactgctggc 8293
QY aagccactaaaataccctgctgaaatctccccactagaagtggtagacacacacagcaagc 8157
Db 8294 aagccactaaaataccctgctgaaatctccccactagaagtggtagacacacacagcaagc 8353
QY acaaaagggcatctcaggacacgtgtgcagaaggtacaagttaaaagagggcctctcagca 8217
Db 8354 acaaaagggcatctcaggacacgtgtgcagaaggtacaagttaaaagagggcctctcagca 8413
QY gtaagtctac 8277
Db 8414 gtaagtctac 8473
QY gataaagggcatcaaacattgaaggaaatctgcataaagacagacacggctccagcagcaagt 8337
Db 8474 gataaagggcatcaaacattgaaggaaatctgcataaagacagacacggctccagcagcaagt 8533
QY gtaactggcagcaggagagcggcccaagcagcaccaggaaggtgcccaagcctagaaagac 8397
Db 8534 gtaactggcagcaggagagcggcccaagcagcaccaggaaggtgcccaagcctagaaagac 8593
QY ctagctggcttcaaaagacccagcagcaggttcacactgaagaaatcaatgactgatacaciaa 8457
Db 8594 ctagctggcttcaaaagacccagcagcaggttcacactgaagaaatcaatgactgatacaciaa 8653
QY accactaaaataccctgcataatcaccagaactagaagacacccgcaacaagctcaag 8517
Db 8654 accactaaaataccctgcataatcaccagaactagaagacacccgcaacaagctcaag 8713
QY agacggcccgagcagcagctgccagaaagttagaagtgaaggagagctgttagcagttggc 8577
Db 8714 agacggcccgagcagcagctgccagaaagttagaagtgaaggagagctgttagcagttggc 8773
QY aagctcacacaaaacctcagggggagacacacacacacacacacacacacacacacacacac 8637
Db 8774 aagctcacacaaaacctcagggggagacacacacacacacacacacacacacacacacacac 8833
QY aagggcacgaaaagcattttaagcaacctgcaaaagcagcgtggacacagagatgtaatt 8697
Db 8834 aagggcacgaaaagcattttaagcaacctgcaaaagcagcgtggacacagagatgtaatt 8893

QY 8698 ggacgagagacagcccaagagcacctaaggaaaaaggcccaacccctggaagacctggcc 8757
|||||
Db 8894 ggacgagagacagcccaagagcacctaaggaaaaaggcccaacccctggaagacctggcc 8953
QY 8758 agcttccaaagagctctctcaaacacacagccacactgaggaaactggcaaatggtctgct 8817
|||||
Db 8954 agcttccaaagagctctctcaaacacacagccacactgaggaaactggcaaatggtctgct 9013
QY 8818 gatagttttaacagcgctcccaagcaaacacactgacagtggaaacctctaaaaatatcc 8877
|||||
Db 9014 gatagttttaacagcgctcccaagcaaacacactgacagtggaaacctctaaaaatatcc 9073
QY 8878 agaagagtcttcgggcccctaaagttagaacccgtgggagacgtggtgaagcaccagagac 8937
|||||
Db 9074 agaagagtcttcgggcccctaaagttagaacccgtgggagacgtggtgaagcaccagagac 9133
QY 8938 cctgtaaaatcacaaagcaaaagcaaacacttccctgcccccaactgccccttaagaaggga 8997
|||||
Db 9134 cctgtaaaatcacaaagcaaaagcaaacacttccctgcccccaactgccccttaagaaggga 9193
QY 8998 ggtggcaaatggaagcgctcacgggaaccaagagcgctgcgtgcacgccagcaccagag 9057
|||||
Db 9194 ggtggcaaatggaagcgctcacgggaaccaagagcgctgcgtgcacgccagcaccagag 9253
QY 9058 gaaattgtggaggagctgccagccagcaagaagcagaggggtgtctcccaggggcaagagcg 9117
|||||
Db 9254 gaaattgtggaggagctgccagccagcaagaagcagaggggtgtctcccaggggcaagagcg 9313
QY 9118 aaatcacccgaacccgtggtcatcatgaagagaaatttgaggactctgcataaaagaatt 9177
|||||
Db 9314 aaatcacccgaacccgtggtcatcatgaagagaaatttgaggactctgcataaaagaatt 9373
QY 9178 gaacctgcggaagagctgaacagcaacgacatgaaacccaacaaagagagaaacacaaatta 9237
|||||
Db 9374 gaacctgcggaagagctgaacagcaacgacatgaaacccaacaaagagagaaacacaaatta 9433
QY 9238 caagactcgctccctgaaataaagggaatatccctgcgtccagcgcgaagataaagact 9297
|||||
Db 9434 caagactcgctccctgaaataaagggaatatccctgcgtccagcgcgaagataaagact 9493
QY 9298 gaggcagaacagcaataactgaggtcttcttattagcagaaaaagaataaataaacaga 9357
|||||
Db 9494 gaggcagaacagcaataactgaggtcttcttattagcagaaaaagaataaataaacaga 9553
QY 9358 aatgaaaaagaccctgaagacctcccagagatggacattcagaatccagatgatga 9417
|||||
Db 9554 aatgaaaaagaccctgaagacctcccagagatggacattcagaatccagatgatga 9613
QY 9418 gcccggaaccccatcacctagagacaaagtctactgagaaacaaaggctgtgaggtctgct 9477
|||||
Db 9614 gcccggaaccccatcacctagagacaaagtctactgagaaacaaaggctgtgaggtctgct 9673
QY 9478 agacagaatgagagctcccgcttaaggtggcagaggagagcgagggcagaaagagtgcg 9537
Db 9674 agacagaatgagagctcccgcttaaggtggcagaggagagcgagggcagaaagagtgcg 9733
QY 9538 aaggttctcatgcagaatcagaaagggaagagagagcaggaattcagactcccatgtgc 9597
|||||
Db 9734 aaggttctcatgcagaatcagaaagggaagagagagcaggaattcagactcccatgtgc 9793
QY 9598 ctgagatcaagaaacacaaaagccagctgcagcaagcactttggagagcaaatctgtg 9657
Db 9794 ctgagatcaagaaacacaaaagccagctgcagcaagcactttggagagcaaatctgtg 9853
QY 9658 cagaagtaaacgcgagtgctcaagggtgtgcagaaaaatccaaaagaagcgtgagacaat 9717
|||||
Db 9854 cagaagtaaacgcgagtgctcaagggtgtgcagaaaaatccaaaagaagcgtgagacaat 9913
QY 9718 gtgtgtgtcaagaaataaacaccagaagctcatagggacagtgaagata 9766
|||||
Db 9914 gtgtgtgtcaagaaataaacaccagaagctcatagggacagtgaagata 9962

RESULT 3
AAA64659
ID AAA64659 standard; cDNA; 11435 BP.
XX
AC AAA64659;
XX
DT 02-JAN-2001 (first entry)
XX
cDNA sequence encoding the antigen of monoclonal antibody Ki-67.
DE
XX Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW antibody Ki-67; ss.
XX
OS Homo sapiens.
XX
PN WO200050595-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04929.
XX
PR 25-FEB-1999; 99US-0257417.
XX
(GOUT/) GOUT I.
PA (RODN/) RODIN N.
PA (FILO/) FILONENKO V.
PA (MATS/) MATSUKA G.
PA (SCAN/) SCANLAN M.
PA (OLDL/) OLD L.
PA (BILY/) BILYNSKY B.
XX
PI Gout I. Rodnin N. Filonenko V. Matsuka G. Scanlan M. Old L;
PI Bilynsky B;
XX
DR WPI: 2000-572092/53.
XX
PT Novel isolated nucleic acid molecules for diagnosing and treating
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -
XX
PS Claim 17; Page 81-85; 94pp; English.
XX
CC The specification describes polynucleotides which are associated with
CC melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC diagnosing and treating a patient with melanoma, thyroid tumour,
CC rectal cancer, lung cancer, breast cancer or colon cancer. The
CC present sequence represents a polynucleotide of the invention.
XX
SQ Sequence 11435 BP; 3802 A; 2828 C; 2684 G; 2121 T; 0 other;

Query Match 84.8%; Score 8286; DB 21; Length 11435;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 8286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1481 atgagagtgaaggaatacccttgaaagagcggtgtgtcccttggcgccactaagac 1540
|||||
Db 597 atgagagtgaaggaatacccttgaaagagcggtgtgtcccttggcgccactaagac 656
QY 1541 ctgaactatttgataaaacttgcctcttaacacgctctcaaaaggagagcccccac 1600
|||||
Db 657 ctgaactatttgataaaacttgcctcttaacacgctctcaaaaggagagcccccac 716
QY 1601 ccaaaagaagtctctggttaatgcacactccacctctctctgaagaataatcaacaagac 1660
|||||
Db 717 ccaaaagaagtctctggttaatgcacactccacctctctctgaagaataatcaacaagac 776
QY 1661 agcctcaaccatcaggaaaaacaagaagctcaggttcagaataatccatgtggaagtgaagcac 1720
|||||
Db 777 agcctcaaccatcaggaaaaacaagaagctcaggttcagaataatccatgtggaagtgaagcac 836

Qy	1721	aaagcttggttataaagccctccagctcctcagtcctcagtaggaaacacccagttgcccagtgatc	1780
Db	837	aaagcttggttataaagccctccagctcctcagtcctcagtaggaaacacccagttgcccagtgatc	896
Qy	1781	aacgcccataagttcctgcacaaacagccccctgcttccagcagcaaatctcagacagaggttc	1840
Db	897	aacgcccataagttcctgcacaaacagccccctgcttccagcagcaaatctcagacagaggttc	956
Qy	1841	ctaagagagggagagaaagatggcaacctgcttcacaaagagagtgctctatcagcgcgaa	1900
Db	957	ctaagagagggagagaaagatggcaacctgcttcacaaagagagtgctctatcagcgcgaa	1016
Qy	1901	gtcaacatgatattttacagatgatattgttcacaaagagaagtggcttcgcgaagacaa	1960
Db	1017	gtcaacatgatattttacagatgatattgttcacaaagagaagtggcttcgcgaagacaa	1076
Qy	1961	atctgattgtgtcaaaatcatgggcagatgtagtaaaacttgggtgcacaaacacacaaa	2020
Db	1077	atctgattgtgtcaaaatcatgggcagatgtagtaaaacttgggtgcacaaacacacaaa	1136
Qy	2021	ctaaagtctataaacacttgcttcctcaaaagtccaatgaacaaagcgcaagaagaccttgcta	2080
Db	1137	ctaaagtctataaacacttgcttcctcaaaagtccaatgaacaaagcgcaagaagaccttgcta	1196
Qy	2081	ctccaaagaagcctgtggcgaaagttccagtcacatcttagtacaggcgccacgcaaacctctc	2140
Db	1197	ctccaaagaagcctgtggcgaaagttccagtcacatcttagtacaggcgccacgcaaacctctc	1256
Qy	2141	cttgtaccataataataggggaaagctcactactgaaaaagtcacatgtgcctgctcgaccct	2200
Db	1257	cttgtaccataataataggggaaagctcactactgaaaaagtcacatgtgcctgctcgaccct	1316
Qy	2201	acagagtgtctcaacaacttcatttccaacccaaataatggactttaaaggagatcttcag	2260
Db	1317	acagagtgtctcaacaacttcatttccaacccaaataatggactttaaaggagatcttcag	1376
Qy	2261	gaatagctgaattgttccaagccccagtgaaaggagcaaccgcagcttgacaagcacatgtc	2320
Db	1377	gaatagctgaattgttccaagccccagtgaaaggagcaaccgcagcttgacaagcacatgtc	1436
Qy	2321	acatcgtctatttcaaatccagagaatttgcttgaaaaacagtttcaaggaaactgattccag	2380
Db	1437	acatcgtctatttcaaatccagagaatttgcttgaaaaacagtttcaaggaaactgattccag	1496
Qy	2381	gagaaagacctctgtccccacctcagagagttttggaggaatgtgtcttccagtgcac	2440
Db	1497	gagaaagacctctgtccccacctcagagagttttggaggaatgtgtcttccagtgcac	1556
Qy	2441	agaatgcagcaaacagccatctgataaatgctctgcagccccctcccttaagacgcgagt	2500
Db	1557	agaatgcagcaaacagccatctgataaatgctctgcagccccctcccttaagacgcgagt	1616
Qy	2501	gtattagaaaaatgaaacgttagcaaaaaccccgaggaacacctacaaaatgacttctc	2560
Db	1617	gtattagaaaaatgaaacgttagcaaaaaccccgaggaacacctacaaaatgacttctc	1676
Qy	2561	tggagacaaaacttcagatctcgacagagccttccaaaacagtatccactgataaaca	2620
Db	1677	tggagacaaaacttcagatctcgacagagccttccaaaacagtatccactgataaaca	1736
Qy	2621	ggtcaggaaaggtctacagagttcaggaatatcacagaagctaccctgttgaaagtaagtg	2680
Db	1737	ggtcaggaaaggtctacagagttcaggaatatcacagaagctaccctgttgaaagtaagtg	1796
Qy	2681	aagaaacaaatcacagaaattgtttgagtgcatactctaaaaagggttcagaaagcaactac	2740
Db	1797	aagaaacaaatcacagaaattgtttgagtgcatactctaaaaagggttcagaaagcaactac	1856
Qy	2741	tacacaaaggagagagagagatgaaaggaataagaaagaccttttgcacacataaagg	2800
Db	1857	tacacaaaggagagagagagatgaaaggaataagaaagaccttttgcacacataaagg	1916
Qy	2801	aaaatttgtaattcaaaagaaaacagatgaaaagatgaaagcaatgaagagatcaagaactt	2860

[illegible]

Db 2997 aagccatgcacacgcctctaaaccatcagtagtgtagagaaagacatcatctatttggg 3056
Qy 3941 gaactccagtgacaaaactggacctgacagagaacttaaccggcgacgaagacggccac 4000
Db 3057 gaactccagtgacaaaactggacctgacagagaacttaaccggcgacgaagacggccac 3116
Qy 4001 aaactcctaagaaagagcccgagctctggaagacctgactggctttaaagagctctcc 4060
Db 3117 aaactcctaagaaagagcccgagctctggaagacctgactggctttaaagagctctcc 3176
Qy 4061 agacccttgcatactagaagaacagtggtgctggtgcacaaactactaaaatgcctcgg 4120
Db 3177 agacccttgcatactagaagaacagtggtgctggtgcacaaactactaaaatgcctcgg 3236
Qy 4121 aatcttctccaccagaaatcagcagacaccccaacaagcacagaagcagcggccaaagac 4180
Db 3237 aatcttctccaccagaaatcagcagacaccccaacaagcacagaagcagcggccaaagac 3296
Qy 4181 ctttgagaaaaagggacctacagaaggagctctcagccctgaagaagctcacacagacat 4240
Db 3297 ctttgagaaaaagggacctacagaaggagctctcagccctgaagaagctcacacagacat 3356
Qy 4241 caggggaaccacacacacagataaaagtaccagaggtgaggataaaagcatcaacgcgt 4300
Db 3357 caggggaaccacacacacagataaaagtaccagaggtgaggataaaagcatcaacgcgt 3416
Qy 4301 ttagggaactgcacaaacagaaaactggaccagcagcaagtgttaactggtagcaagagcc 4360
Db 3417 ttagggaactgcacaaacagaaaactggaccagcagcaagtgttaactggtagcaagagcc 3476
Qy 4361 acccaaaaacttaagaaaaagggcccaaccccttagaagacctggctggctggaagagctct 4420
Db 3477 acccaaaaacttaagaaaaagggcccaaccccttagaagacctggctggctggaagagctct 3536
Qy 4421 tccagaccagtgatgactacagaagccacagactccagagaaactaccaaataagct 4480
Db 3537 tccagaccagtgatgactacagaagcccaagcctcagagaaactaccaaataagct 3596
Qy 4481 gcagatcaaacaccagaccagtgacacacccaacaagctccaagccacagctccaagaaa 4540
Db 3597 gcagatcaaacaccagaccagtgacacacccaacaagctccaagccacagctccaagaaa 3656
Qy 4541 gtctcagaaaagtgagactagaagaagaattcttcgcactcaggaagaaacacacacatcag 4600
Db 3657 gtctcagaaaagtgagactagaagaagaattcttcgcactcaggaagaaacacacacatcag 3716
Qy 4601 caggcaaaagccatgcacacacccaacacagcagtaagtggtagaataaacatctacgcat 4660
Db 3717 caggcaaaagccatgcacacacccaacacagcagtaagtggtagaataaacatctacgcat 3776
Qy 4661 ttatgggaactccagtcagaaaactggacctgacagagaacttaactggcgcaagagac 4720
Db 3777 ttatgggaactccagtcagaaaactggacctgacagagaacttaactggcgcaagagac 3836
Qy 4721 ggctacaaaactcctaagaaaagggcccgagctcttagaagacctggctggctttaaagagc 4780
Db 3837 ggctacaaaactcctaagaaaagggcccgagctcttagaagacctggctggctttaaagagc 3896
Qy 4781 tcttcagacacagaggtcacactaggaataaatgactaaacgataaaaactcccaagtag 4840
Db 3897 tcttcagacacagaggtcacactaggaataaatgactaaacgataaaaactcccaagtag 3956
Qy 4841 cctgcaaatcttcaaacaccagaccttagacaaaaccccgcaagctcccaagcgagcggtca 4900
Db 3957 cctgcaaatcttcaaacaccagaccttagacaaaaccccgcaagctcccaagcgagcggtca 4016
Qy 4901 agacatccctgggaaaaagtgggctgtaaaagaagagctccttagcagttggcaagctcacac 4960
Db 4017 agacatccctgggaaaaagtgggctgtaaaagaagagctccttagcagttggcaagctcacac 4076
Qy 4961 agacatcaggagagactcacacacacacacacagagcccaacagggagatggttaagcatga 5020
Db 4077 agacatcaggagagactcacacacacacacacagagcccaacagggagatggttaagcatga 4136

Qy 5021 aagcatttatgagtgctccaaagcagatcttagactcagcagaacttaacttggcagca 5080
Db 4137 aagcatttatgagtgctccaaagcagatcttagactcagcagaacttaacttggcagca 4196
Qy 5081 agaggcagctgagaactccttaaggaaaagctgaaagtcctgaagtcctgaagacctggcggttca 5140
Db 4197 agaggcagctgagaactccttaaggaaaagctgaaagtcctgaagacctggcggttca 4256
Qy 5141 tggagcttccagacacacaaactcacactaaggaataatgaataacttaaaaaactacca 5200
Db 4257 tggagcttccagacacacaaactcacactaaggaataatgaataacttaaaaaactacca 4316
Qy 5201 aagttactacagagcttcacagccagacctagtgacaccccaacaagaactccaagccac 5260
Db 4317 aagttactacagagcttcacagccagacctagtgacaccccaacaagaactccaagccac 4376
Qy 5261 agcccaagagaagctctcaggaaaagcagacactgaaagaagaatttttagcatttaggaac 5320
Db 4377 agcccaagagaagctctcaggaaaagcagacactgaaagaagaatttttagcatttaggaac 4436
Qy 5321 aaacgccatcagcagcaaaagccatgcacacaccccaaccacccagcagtagtgaaagaaag 5380
Db 4437 aaacgccatcagcagcaaaagccatgcacacaccccaaccacccagcagtagtgaaagaaag 4496
Qy 5381 acatacaacagtttttgggaactccagtcgagaaaactggaccagccaggaataatttacctg 5440
Db 4497 acatacaacagtttttgggaactccagtcgagaaaactggaccagccaggaataatttacctg 4556
Qy 5441 gcagcaatagacggtctcaaaactcgtaaaggaaaagcccaggtccttagaagacctgactg 5500
Db 4557 gcagcaatagacggtctcaaaactcgtaaaggaaaagcccaggtccttagaagacctgactg 4616
Qy 5501 gcttcagagagcttttcagacacaccatgcactgataaacccacagctgatagaataacta 5560
Db 4617 gcttcagagagcttttcagacacaccatgcactgataaacccacagctgatagaataacta 4676
Qy 5561 ccaaaaaatactctgcaaatctccgcaatcagacccagcggacaccccaacaacacaa 5620
Db 4677 ccaaaaaatactctgcaaatctccgcaatcagacccagcggacaccccaacaacacaa 4736
Qy 5621 agcaacgcccgaagagaagcctcaagaaaagcagtagagaagaagaatttttagcattca 5680
Db 4737 agcaacgcccgaagagaagcctcaagaaaagcagtagagaagaagaatttttagcattca 4796
Qy 5681 ggaactaacacatcagcagcgaagccatgcacacgcctaaagcagcagtagtgagag 5740
Db 4797 ggaactaacacatcagcagcgaagccatgcacacgcctaaagcagcagtagtgagag 4856
Qy 5741 ggaagacatcaacacatttgggactccagtgagaaaactggacctgtaggaatt 5800
Db 4857 ggaagacatcaacacatttgggactccagtgagaaaactggacctgtaggaatt 4916
Qy 5801 tacttggcagaagagaagccacaaaactccttaagaaaaagccaggctctagaagatc 5860
Db 4917 tacttggcagaagagaagccacaaaactccttaagaaaaagccaggctctagaagatc 4976
Qy 5861 tggctggcttcaaaagagctcttcagacacacaggtcacactgagaatacaatgacgcag 5920
Db 4977 tggctggcttcaaaagagctcttcagacacacaggtcacactgagaatacaatgacgcag 5036
Qy 5921 acaaaatcacagaagtaactctgcaaatctcccaaccagacccagtcacaaaccccaacaa 5980
Db 5037 acaaaatcacagaagtaactctgcaaatctcccaaccagacccagtcacaaaccccaacaa 5096
Qy 5981 gctccaagcaacgactcaagatatccttgggaaaagtagtgtgaagaagagggtcctac 6040
Db 5097 gctccaagcaacgactcaagatatccttgggaaaagtagtgtgaagaagagggtcctac 5156
Qy 6041 cagtcggaagctcacacagcgtcagggaagacacacacagacacagacagagacagag 6100
Db 5157 cagtcggaagctcacacagcgtcagggaagacacacacagacacagacagagacagag 5216

QY	6101	gagatgaaagagcatcaaacgctttaaggaattctgcaaaagcagatgctggaacccagcaa	6160
Db	5217	gagatgaaagagcatcaaaagcgtttaaggaattctgcaaaagcagatgctggaacccagcaa	5276
QY	6161	actatggaactggatggagagtggtgccaagaacacctaaagagggcccaatcactag	6220
Db	5277	actatggaactggatggagagtggtgccaagaacacctaaagagggcccaatcactag	5336
QY	6221	aagacctggccggttcaaaagagctctccagacacacagaccactgaggaatacaaca	6280
Db	5337	aagacctggccggttcaaaagagctctccagacacacagaccactgaggaatacaaca	5396
QY	6281	ctgatgacaaaactaccaaaatagctgcaaaatctccaccaccagaaatcaatggacactc	6340
Db	5397	ctgatgacaaaactaccaaaatagctgcaaaatctccaccaccagaaatcaatggacactc	5456
QY	6341	caacaagcacaaaggcgcgccaaaacacctttggggaagggaatatagtggaagagc	6400
Db	5457	caacaagcacaaaggcgcgccaaaacacctttggggaagggaatatagtggaagagc	5516
QY	6401	tctcagccctgaagcagctcacacagaccacacacagacaaagtaccagggagatgagg	6460
Db	5517	tctcagccctgaagcagctcacacagaccacacacagacaaagtaccagggagatgagg	5576
QY	6461	ataaaggatcaacgctgttcagggaactgcataaaacagaaactggacccagcagcaagt	6520
Db	5577	ataaaggatcaacgctgttcagggaactgcataaaacagaaactggacccagcagcaagt	5636
QY	6521	taactggttagcagagcagcgaagaactcctaaagggaagcccaacccctagaagact	6580
Db	5637	taactggttagcagagcagcgaagaactcctaaagggaagcccaacccctagaagact	5696
QY	6581	tggctggcttgaagagctcttcagacaccagtagtgcactgcaagcccaacactcaag	6640
Db	5697	tggctggcttgaagagctcttcagacaccagtagtgcactgcaagcccaacactcaag	5756
QY	6641	agaaaactaccaaaatagcctgcagatctccacaacacagaccagtcgggtaccaccaaca	6700
Db	5757	agaaaactaccaaaatagcctgcagatctccacaacacagaccagtcgggtaccaccaaca	5816
QY	6701	tcttcaagccacagtccaagagcttcagggaagcagacagctagaggaagaatccttag	6760
Db	5817	tcttcaagccacagtccaagagcttcagggaagcagacagctagaggaagaatccttag	5876
QY	6761	cactcaggaaacgaacacacatcagtagtggaagcctatggacacaccccaacccagcagg	6820
Db	5877	cactcaggaaacgaacacacatcagtagtggaagcctatggacacaccccaacccagcagg	5936
QY	6821	gtgatgagaagacatgaaaagcatttatgggaactccagtcgagaaattggacctgccag	6880
Db	5937	gtgatgagaagacatgaaaagcatttatgggaactccagtcgagaaattggacctgccag	5996
QY	6881	gaatttacctggcagcaaaaagatggcccaaaactcctaaaggaaaggccaggtccttag	6940
Db	5997	gaatttacctggcagcaaaaagatggcccaaaactcctaaaggaaaggccaggtccttag	6056
QY	6941	aagacctggcttcaaaagagctcttcagacacacagcagcactgcaagcccaacgactg	7000
Db	6057	aagacctggcttcaaaagagctcttcagacacacagcagcactgcaagcccaacgactg	6116
QY	7001	atgagaaaactaccaaaatagcctgcaaaatctccacaacccagaccagtcggaaccccaag	7060
Db	6117	atgagaaaactaccaaaatagcctgcaaaatctccacaacccagaccagtcggaaccccaag	6176
QY	7061	caagcaaaagcaacggcccaagagaacctcaggaagcagacgtagaggaagaatttt	7120
Db	6177	caagcaaaagcaacggcccaagagaacctcaggaagcagacgtagaggaagaatttt	6236
QY	7121	tagcactcaggaaacgaacacatcagcaggcgaagccatggacaccccaacccagcag	7180
Db	6237	tagcactcaggaaacgaacacatcagcaggcgaagccatggacaccccaacccagcag	6296
QY	7181	taagtgatgagaaaaatatcaacacatttgtggaactccagtcgagaaactggacctgc	7240

Db	6297	taagtgatgagaaaaatatcaacacatttgtggaactccagtcgagaactggacctgc	6356
QY	7241	taggaatttaccctggcagcaagagagacagaccagactcctaaagaaaagcctgaggtcc	7300
Db	6357	taggaatttaccctggcagcaagagagacagaccagactcctaaagaaaagcctgaggtcc	6416
QY	7301	tagagacctggcttcaaaagaaactctccagacacccaggttcacactgaggaatacaa	7360
Db	6417	tagagacctggcttcaaaagaaactctccagacacccaggttcacactgaggaatacaa	6476
QY	7361	tgactgatgacaaaaatcaacagaagtatcctgttaaatctccacagcagagtcatctcaaaa	7420
Db	6477	tgactgatgacaaaaatcaacagaagtatcctgttaaatctccacagcagagtcatctcaaaa	6536
QY	7421	cctcaagaagctccaagcaaaagcctcaagatccccctggtgaaagtggacatgaagaag	7480
Db	6537	cctcaagaagctccaagcaaaagcctcaagatccccctggtgaaagtggacatgaagaag	6596
QY	7481	agccctagcagtcagcaagctcacacgggacatcaggggagactagcgaaacacacacag	7540
Db	6597	agccctagcagtcagcaagctcacacgggacatcaggggagactagcgaaacacacacag	6656
QY	7541	agccacagagagatagtaagagacatcaaaagcgtttaaggaggtctccaaagcagatcctgg	7600
Db	6657	agccacagagagatagtaagagacatcaaaagcgtttaaggaggtctccaaagcagatcctgg	6716
QY	7601	accacagcaagcttaactggttagcagagcagctgagaactcgttaagaaaagggccc	7660
Db	6717	accacagcaagcttaactggttagcagagcagctgagaactcgttaagaaaagggccc	6776
QY	7661	gtgctctagaagacctgggtgacttcaaaagagctcttctcagcaccaggttcacactgaag	7720
Db	6777	gtgctctagaagacctgggtgacttcaaaagagctcttctcagcaccaggttcacactgaag	6836
QY	7721	agtcaatgactattgcataaaacacaaaaatccccctgcaaatccccccaccagaaactaa	7780
Db	6837	agtcaatgactattgcataaaacacaaaaatccccctgcaaatccccccaccagaaactaa	6896
QY	7781	cagacactgccagcagcaaaagagatgccccagacagctcccaggaagagatataaag	7840
Db	6897	cagacactgccagcagcaaaagagatgccccagacagctcccaggaagagatataaag	6956
QY	7841	aggagctctcagcagttgagaggtctcagcaaaacatcagggcaaaagcacacacacacaca	7900
Db	6957	aggagctctcagcagttgagaggtctcagcaaaacatcagggcaaaagcacacacacacaca	7016
QY	7901	aagaaaccagcaagcgtgtatgagggcatcaaaagtatttgaaagcaacgtgcaaaagaaaac	7960
Db	7017	aagaaaccagcaagcgtgtatgagggcatcaaaagtatttgaaagcaacgtgcaaaagaaaac	7076
QY	7961	caaaaccagtagaagagggaacccagcaggaagggccaagagcaccttaagggaagaaagccc	8020
Db	7077	caaaaccagtagaagagggaacccagcaggaagggccaagagcaccttaagggaagaaagccc	7136
QY	8021	aacccctggaagacctggccggttccacagagctctctgaaacatcagggttcacactcagg	8080
Db	7137	aacccctggaagacctggccggttccacagagctctctgaaacatcagggttcacactcagg	7196
QY	8081	aatcactgactgctggcaaaagcccaactaaaaataacctgcaaatccccccactagaagtgg	8140
Db	7197	aatcactgactgctggcaaaagcccaactaaaaataacctgcaaatccccccactagaagtgg	7256
QY	8141	tagacaccacagcaacacaaagagcattctcaggacacgctgaggaaggttacaagtaa	8200
Db	7257	tagacaccacagcaacacaaagagcattctcaggacacgctgaggaaggttacaagtaa	7316
QY	8201	aagaaagaccttcagcagtcgaagtccacacaaacatcagggggaacccaggaatgcagaca	8260
Db	7317	aagaaagaccttcagcagtcgaagtccacacaaacatcagggggaacccaggaatgcagaca	7376
QY	8261	aagaaaccagcaggtgaagataaaaggatcaaaagcattgaaaggaattctgcaaaacagacac	8320

Db 7377 aagaaccagcggtgaagataaaggcatcaagcattgaaggaatctctcaaaacagacac 7436
Qy 8321 cggctccagcagcaagttaactggcagcagagagacgcccgaagacacccagggaagtg 8380
Db 7437 cggctccagcagcaagttaactggcagcagagagacgcccgaagacacccagggaagtg 7496
Qy 8381 cccaagccatagaagacactagctggcttcaagacccccagcagcagggtcacactgaagaat 8440
Db 7497 cccaagccatagaagacactagctggcttcaagacccccagcagcagggtcacactgaagaat 7556
Qy 8441 caatgactgatgacaaaaccactaaataaccctgcaaatcatcaccagaactagaaagaca 8500
Db 7557 caatgactgatgacaaaaccactaaataaccctgcaaatcatcaccagaactagaaagaca 7616
Qy 8501 ccgcaacaagctcgaagagacgcccagagacacgtgcccgaagaagtgaagtgaaggagg 8560
Db 7617 ccgcaacaagctcgaagagacgcccagagacacgtgcccgaagaagtgaagtgaaggagg 7676
Qy 8561 agctgttagcagttggcgaagctcacacaaacctcaggggagaccacgcacaccgcacaag 8620
Db 7677 agctgttagcagttggcgaagctcacacaaacctcaggggagaccacgcacaccgcacaag 7736
Qy 8621 agccggtaggtgagggcgaagagcagcaaaagcatttaagcaacctgcgaagcggaacgtgg 8680
Db 7737 agccggtaggtgagggcgaagagcagcaaaagcatttaagcaacctgcgaagcggaacgtgg 7796
Qy 8681 acgcagaagatgtaattggcagcagagacagcgaacacctaagagcactaaaggaaagcccaac 8740
Db 7797 acgcagaagatgtaattggcagcagagacagcgaacacctaagagcactaaaggaaagcccaac 7856
Qy 8741 cccctggaagacctggccagcttccaagagctctccaagagctctccaacaccaggccacactgagggaac 8800
Db 7857 cccctggaagacctggccagcttccaagagctctccaacaccaggccacactgagggaac 7916
Qy 8801 tggcaaatggtgctgctatagctttacaagcgtccaaagcgaacacactacacactgagaa 8860
Db 7917 tggcaaatggtgctgctatagctttacaagcgtccaaagcgaacacactacacactgagaa 7976
Qy 8861 aacctctaaaaatccagaagagttcttcgggcccctaaagttagaaccctgaggagagc 8920
Db 7977 aacctctaaaaatccagaagagttcttcgggcccctaaagttagaaccctgaggagagc 8036
Qy 8921 tggtaagcaccagagacctgtaaaaatcacaaagcgaacaaacttccctgccccccac 8980
Db 8037 tggtaagcaccagagacctgtaaaaatcacaaagcgaacaaacttccctgccccccac 8096
Qy 8981 tgccttcaagaggggaggtggcaaaagtgaagcgtcacgggaaccaggaggtgcgct 9040
Db 8097 tgccttcaagaggggaggtggcaaaagtgaagcgtcacgggaaccaggaggtgcgct 8156
Qy 9041 gcatgccagccaccagagaaattgtggagagctgccagccagcaagacagagaggttg 9100
Db 8157 gcatgccagccaccagagaaattgtggagagctgccagccagcaagacagagaggttg 8216
Qy 9101 ctcccagggcaagaggcaaaatcattccgaacccctggtgctcatcatgaagagaagtttgagga 9160
Db 8217 ctcccagggcaagaggcaaaatcattccgaacccctggtgctcatcatgaagagaagtttgagga 8276
Qy 9161 ctcttgcataaagaattgaacctgcggaagagctgaacgaacagacatgataaaaccaaca 9220
Db 8277 ctcttgcataaagaattgaacctgcggaagagctgaacgaacagacatgataaaaccaaca 8336
Qy 9221 aaggggaacacaaattacgaactcgttccctgaaaaataagggaatatccctggcctcca 9280
Db 8337 aaggggaacacaaattacgaactcgttccctgaaaaataagggaatatccctggcctcca 8396
Qy 9281 gacgccaagataaagactgagcagaacacagcaataactagtgaggtctttgtattagcagaaa 9340
Db 8397 gacgccaagataaagactgagcagaacacagcaataactagtgaggtctttgtattagcagaaa 8456
Qy 9341 gaatagaataaacaagaataaagaagcccatgaagacacctccccagagatggacattc 9400
Db 8457 gaatagaataaacaagaataaagaagcccatgaagacacctccccagagatggacattc 8516

Qy 9401 agaattccagatgatgagcccgaaaaccatacctagagacaaagtctactgagaacaaaa 9460
Db 8517 agaattccagatgatgagcccgaaaaccatacctagagacaaagtctactgagaacaaaa 8576
Qy 9461 ggtgcttaggtctctctagacagaatagagagctccagcctaaagtggcagaggagagcg 9520
Db 8577 ggtgcttaggtctctctagacagaatagagagctccagcctaaagtggcagaggagagcg 8636
Qy 9521 gaggcacaagagtgcgaaagttctcatgcacaatcagaagaaggaaagagagcaggaa 9580
Db 8637 gaggcacaagagtgcgaaagttctcatgcagaaacagaaaaggaaagagcaggaa 8696
Qy 9581 attcagactccatgtgcttgagatcaagaagacaaaaagccagcctgcagcaagcactt 9640
Db 8697 attcagactccatgtgcttgagatcaagaagacaaaaagccagcctgcagcaagcactt 8756
Qy 9641 tggagagaaaattctgtgcagagagtaacgcggagtgctcaagagggtgcagaaaaatccaa 9700
Db 8757 tggagagaaaattctgtgcagagagtaacgcggagtgctcaagagggtgcagaaaaatccaa 8816
Qy 9701 agaagctggaagacaaatgtgtgtcagaagaaaataacaaccagaagtcattagggacagtg 9760
Db 8817 agaagctggaagacaaatgtgtgtcagaagaaaataacaaccagaagtcattagggacagtg 8876
Qy 9761 aagata 9766
Db 8877 aagata 8882

RESULT 4
AAH20873
ID AAH20873 standard; DNA; 733 BP.
XX
AC AAH20873;
XX
DT 21-AUG-2001 (first entry)
XX
DE Human Ki-67 protein KON-21 DNA fragment.
XX
KW Ki-67 protein; human; KON-21; anticancer; anti-allergic; gene therapy; immunomodulatory; anti-inflammatory; anti-rheumatic; cancer; allergy; autoimmune disease; inflammation; rheumatic disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 60..725
FT /*tag= a
FT /product= "KON-21"
XX
PN WO200136629-A1.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-EP11482.
XX
PR 18-NOV-1999; 95DE-1055576.
XX
PA (BORS-) FORSCHUNGSZENTRUM BORSTEL.
XX
PI Gerdes J, Scholzen T, Wohlenberg C;
XX
DR WPI; 2001-355634/37.
XX
DR P-PSDB; AAB86191.
XX
PT System for transferring compounds between cells, useful e.g. for delivery of drugs or gene therapy sequences, based on the C-terminal fragment of human protein Ki-67 -
XX
PS Disclosure; Fig 1; 3lpp; German.
XX
CC This invention describes a novel use of the C-terminal fragment (A) of

CC Ki-67 protein (or its parts, fragments or homologs) for transfer of
CC attached compounds (B) into cells, or for uptake of (B) by, or release of
CC (B) from cells. The Ki-67 C-terminal fragment, designated KON-21, has
CC anticancer, anti-allergic, immunomodulatory, anti-inflammatory and
CC anti-rheumatic activity. (A), or transfer proteins (I) containing (A) and
CC at least one other component, are used as carriers for other active
CC ingredients in pharmaceutical compositions and to produce gene therapy
CC compositions, for treatment or prevention of, specifically, cancer,
CC allergy, autoimmune diseases, inflammation and rheumatic diseases. The
CC transfer vehicles are of human origin, avoiding risks associated with
CC using viral carriers. This sequence encodes the KON-21 protein fragment
CC used in the method of the invention.
XX
SQ Sequence 733 BP; 271 A; 144 C; 199 G; 119 T; 0 other;

Query Match 6.4%; Score 629; DB 22; Length 733;
Best Local Similarity 100.0%; Pred. No. 3.6e-284; Indels 0; Gaps 0;
Matches 629; Conservative 0; Mismatches 0;
QY 9108 ggcaagaggcnaatcatccgaaccctgggtcatcatgaagagaagtgtgaggactcttcg 9167
Db 62 ggcaagaggcnaatcatccgaaccctgggtcatcatgaagagaagtgtgaggactctgc 121
QY 9168 aaaaagaattgaacctgcggaaagctgaacagcaacgacatgaaaacccaacaaagagga 9227
Db 122 aaaaagaattgaacctgcggaaagctgaacagcaacgacatgaaaacccaacaaagagga 181
QY 9228 acacaattacagacctgcctcctgaataaaggaaatccctgcgtccagacgcga 9287
Db 182 acacaattacagacctgcctcctgaataaaggaaatccctgcgtccagacgcga 241
QY 9288 agataagactgagcgagaaacagcaaaataactgaggtcttctgtattagcagaaagataga 9347
Db 242 agataagactgagcgagaaacagcaaaataactgaggtcttctgtattagcagaaagataga 301
QY 9348 aataacagaaaatgaaaagaagccctgaagacctccccagagatggaattcagaatcc 9407
Db 302 aataacagaaaatgaaaagaagccctgaagacctccccagagatggaattcagaatcc 361
QY 9408 agatgatgagcccgaaacccatacctagagacaagaagtcactgaacaaaggtgctt 9467
Db 362 agatgatgagcccgaaacccatacctagagacaagaagtcactgaacaaaggtgctt 421
QY 9468 gaggtctgtagacagaatgagagctccagcctaaggtggcagagagcgagggca 9527
Db 422 gaggtctgtagacagaatgagagctccagcctaaggtggcagagagcgagggca 481
QY 9528 gaagagtgcgaaggttctcatgcagaatcagaagaaggaaagagagcgaggaattcaga 9587
Db 482 gaagagtgcgaaggttctcatgcagaatcagaagaaggaaagagagcgaggaattcaga 541
QY 9588 ctccatgtcctgagatcaagaagaacaaaagccagcctgcagcaagcactttgagag 9647
Db 542 ctccatgtcctgagatcaagaagaacaaaagccagcctgcagcaagcactttgagag 601
QY 9648 caaatctgtcagagagtagtaacgcggaggtgtcaagagaggtgtgcagaaaaatccaaagagcg 9707
Db 602 caaatctgtcagagagtagtaacgcggaggtgtcaagagaggtgtgcagaaaaatccaaagagcg 661
QY 9708 tgaggacaatgtgtgtgtcaagaaaaataa 9736
Db 662 tgaggacaatgtgtgtgtcaagaaaaataa 690

RESULT 5
AAA64645
ID AAA64645 standard; DNA; 904 BP.
XX
AC
XX
DT 02-JAN-2001 (first entry)
XX

DE Partial sequence MEL10 of a Ki-67 nuclear antigen gene.
XX
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW Ki-67 nuclear antigen gene; ss.
XX
OS Homo sapiens.
XX
PN WO2000050595-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04929.
XX
PR 25-FEB-1999; 99US-0257417.
XX
PA (GOUT/) GOUT I.
PA (RODN/) RODNIN N.
PA (FILO/) FILOENKO V.
PA (MATS/) MATSUKA G.
PA (SCAN/) SCANLAN M.
PA (OLDL/) OLD L.
PA (BILY/) BILYNSKY B.
XX
PI Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
PI Bilynsky B;
XX
DR WPI; 2000-572092/53.
XX
PT Novel isolated nucleic acid molecules for diagnosing and treating
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -
XX
PS Example 2; Page 66-67; 94pp; English.
XX
CC The specification describes polynucleotides which are associated with
CC melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC diagnosing and treating a patient with melanoma, thyroid tumour,
CC rectal cancer, lung cancer, breast cancer or colon cancer. The
CC present sequence represents a partial sequence of a Ki-67 nuclear
CC antigen gene, and is a polynucleotide of the invention.
XX
SQ Sequence 904 BP; 318 A; 227 C; 221 G; 127 T; 11 other;

Query Match 5.4%; Score 527; DB 21; Length 904;
Best Local Similarity 99.7%; Pred. No. 2.3e-236;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6307 tgcaaatctccaccaccagaatcaatggacactccaacaagcacaaggcgcccaaa 6366
Db 276 tgcaaatctccaccaccagaatcaatggacactccaacaagcacaaggcgcccaaa 335
QY 6367 acacctttgggaaaaaggatagtggaagagctctcagccctgaagcagctcacacag 6426
Db 336 acacctttgggaaaaaggatagtggaagagctctcagccctgaagcagctcacacag 395
QY 6427 accacacacagacaaagtaccaggagatgaggaataaaggcacaacgtgttcaggga 6486
Db 396 accacacacagacaaagtaccaggagatgaggaataaaggcacaacgtgttcaggga 455
QY 6487 actcaaaaaacagaactggaccaccagcagcaagtgttaactggtagcaaggcgccaaga 6546
Db 456 actcaaaaaacagaactggaccaccagcagcaagtgttaactggtagcaaggcgccaaga 515
QY 6547 actcctaagggaaaagcccaacccctagaaagacttggcttggttgaaagagctctccag 6606
Db 516 actcctaagggaaaagcccaacccctagaaagacttggcttggttgaaagagctctccag 575
QY 6607 acaccagatgcacigacaagcccgactcacgagaaataaccacaaatagcctgcaga 6666
Db 576 acaccaataatgcactgacaagcccgcgactcatgagaaaaactaccataagcctgcaga 635

QY 6667 tctccacaaccagaccagtggtgtaccaccaacaatcttcaagccagtcctcaagagaagt 6726
Db 636 tctccacaaccagaccagtggtgtaccaccaacaatcttcaagccagtcctcaagagaagt 695
QY 6727 ctccaggaagcagcgtagagaagaatccttagcactcaggaagaacgaacacccatcagta 6786
Db 696 ctccaggaagcagcgtagagaagaatccttagcactcaggaagaacgaacacccatcagta 755
QY 6787 gggaaagctatggacacaccccaaccagcagagtgatgagaagaacagatgaagcattt 6846
Db 756 gggaaagctatggacacaccccaaccagcagagtgatgagaagaacagatgaagcattt 815
QY 6847 atgggaactcagtcagaaattggacctgcccaggaatttacctgggcagcaaaaagatgg 6906
Db 816 atgggaactcagtcagaaattggacctgcccaggaatttacctgggcagcaaaaagatgg 875
QY 6907 ccacaactccttaaggaagggccagc 6935
Db 876 ccacaactccttaaggaagggccagc 904

RESULT 6
AAZ80374
ID AAZ80374 standard; cDNA; 561 BP.
XX
AC AAZ80374;
XX
DT 07-APR-2000 (first entry)
XX

DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:458.
XX
KW Human; gene expression product; diagnosis; tumour; colon cancer;
KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
KW hyperplasia; ds.
XX

OS Homo sapiens.
XX
PN WO9964576-A2.
XX
PD 16-DEC-1999.
XX

PF 09-JUN-1999; 99WO-IB01062.
XX
PR 10-JUN-1998; 98US-0088801.
XX
PA (FARB) BAYER CORP.
XX

PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
PI Carroll E, Catino TJ, Derti A, Ford DM, Lewis ME, Monahan JE;
PI Schlegel R;
XX
DR WPI; 2000-087220/07.
XX

PT Novel nucleic acids, used to develop products for the diagnosis and
PT treatment of disorders involving unwanted cell proliferation,
PT particularly cancers, especially colon cancer -
XX
PS Claim 15; Page 314; 469pp; English.
XX

CC AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
CC the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC cDNA clones can be used to generate antisense oligonucleotides which
CC can be used for antisense therapy. Methods and products from the present
CC invention can be used for identifying and/or classifying cancerous cells
CC present in a human tumour, particularly in solid tumours, e.g.
CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC can be used for developing agents for the diagnosis and treatment of
CC disorders involving unwanted cell proliferation, such as neoplasia,
CC dysplasia or hyperplasia.
XX

SQ Sequence 561 BP; 196 A; 143 C; 125 G; 81 T; 16 other;

Query Match 4.0%; Score 389; DB 21; Length 561;
Best Local Similarity 100.0%; Pred. No. 1.1e-171;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6691 accccaacaatttcaagccacagtcacagagaagtctcaggaagacagacgtagaggaa 6750
Db 1 accccaacaatttcaagccacagtcacagagaagtctcaggaagacagacgtagaggaa 60
QY 6751 gaatctagcactcaggaacgaacacacatcagtaggaaagctatgacacacccaaa 6810
Db 61 gaatcttagcactcaggaacgaacacacatcagtaggaaagctatgacacacccaaa 120
QY 6811 ccagcaggaggtgatgagaagacatgaaagcattttatggaaactccagtcagaaattg 6870
Db 121 ccagcaggaggtgatgagaagacatgaaagcattttatggaaactccagtcagaaattg 180
QY 6871 gacctgcaagaaatttacctggcagcaaaaagatggcccaaaactcctaaagaaagggcc 6930
Db 181 gacctgcaagaaatttacctggcagcaaaaagatggcccaaaactcctaaagaaagggcc 240
QY 6931 caggctctagaagacctggtgttcaaaagagctcttcagacacacagcactgacaag 6990
Db 241 caggctctagaagacctggtgttcaaaagagctcttcagacacacagcactgacaag 300
QY 6991 cccagcactgatgagaaaaactacaaaatagcctgcaaatctccacaaccagccagtg 7050
Db 301 cccagcactgatgagaaaaactacaaaatagcctgcaaatctccacaaccagccagtg 360
QY 7051 gacaccccagcagcacaaaagcaagcggcc 7079
Db 361 gacaccccagcagcacaaaagcaagcggcc 389

RESULT 7
AA64643
ID AAA64643 standard; DNA; 850 BP.
XX
AC AAA64643;
XX
DT 02-JAN-2001 (first entry)
XX
DE Partial sequence MEL8 of a Ki-67 nuclear antigen gene.
XX
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW Ki-67 nuclear antigen gene; ss.
XX

OS Homo sapiens.
XX
PN WO200050595-A2.
XX
PD 31-AUG-2000.
XX

PF 25-FEB-2000; 2000WO-US04929.
XX
PR 25-FEB-1999; 99US-0257417.
XX
PA (GOUT/) GOUT I.
PA (RODN/) RODNIN N.
PA (FILO/) FILOENKO V.
PA (MATS/) MATSUKA G.
PA (SCAN/) SCANLAN M.
PA (OLDL/) OLD L.
PA (BILY/) BILYNSKY B.
XX

XX Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
PI Bilynsky B;
XX
DR WPI; 2000-572092/53.
XX

PT Novel isolated nucleic acid molecules for diagnosing and treating
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -

XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
XX
XX
SQ

Example 2; Page 65; 94pp; English.

The specification describes polynucleotides which are associated with melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5, Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for diagnosing and treating a patient with melanoma, thyroid tumour, rectal cancer, lung cancer, breast cancer or colon cancer. The present sequence represents a partial sequence of a Ki-67 nuclear antigen gene, and is a polynucleotide of the invention.

Sequence 850 BP; 302 A; 208 C; 197 G; 129 T; 14 other;

Query Match 3.8%; Score 370; DB 21; Length 850;
Best Local Similarity 99.6%; Pred. No. 9,1e-163;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6464 aaggcatcaacgtgttctcagggaactgcacacagaaaactggaccagcagcaagtgttaa 6523
|||||
Db 379 aaggcatcaacgtgttctcagggaactgcacacagaaaactggaccagcagcaagtgttaa 438
QY 6524 ctggtagcagagcagcagcaactccttaagggaagcccaacccttagaagacttg 6583
|||||
Db 439 ctggtagcagagcagcagcaactccttaagggaagcccaacccttagaagacttg 498
QY 6584 ctggttgaagagctcttcacagacacagctatgctgacagcccaagcactcacgaga 6643
|||||
Db 499 ctggttgaagagctcttcacagacacagctatgctgacagcccaagcactcacgaga 558
QY 6644 aaactacaaaatagctgcagatcttcacacacagaccagctgggtaccccaacaatct 6703
|||||
Db 559 aaactacaaaatagctgcagatcttcacacacagaccagctgggtaccccaacaatct 618
QY 6704 tcaagcacagtcacagagagtcctcagagaaagcagacagctagaggaatccttagcac 6763
|||||
Db 619 tcaagcacagtcacagagagtcctcagagaaagcagacagctagaggaatccttagcac 678
QY 6764 tcaggaaacgaacaccatcagtagggaagctatggacacaccccaaccagcaggaggtg 6823
|||||
Db 679 tcaggaaacgaacaccatcagtagggaagctatggacacaccccaaccagcaggaggtg 738
QY 6824 atgagaaagacatgaaagcattttatgggaactccagtcagaaattggacctgccaggaa 6883
|||||
Db 739 atgagaaagacatgaaagcattttatgggaactccagtcagaaattggacctgccaggaa 798
QY 6884 atttacctgcagcaaaaagatgcccaactccttaaggaaaagcccaggc 6935
|||||
Db 799 atttacctgcagcaaaaagatgcccaactccttaaggaaaagcccaggc 850

RESULT 8

AAA64639
ID AAA64639 standard; DNA; 861 BP.

XX
AC AAA64639;

XX
DT 02-JAN-2001 (first entry)

XX
DE Partial sequence MEL4 of a Ki-67 nuclear antigen gene.

XX
KW Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW rectal cancer; lung cancer; breast cancer; colon cancer;
KW Ki-67 nuclear antigen gene; ss.

OS Homo sapiens.

XX
PN WO200050595-A2.

XX
PD 31-AUG-2000.

XX
PF 25-FEB-2000; 2000WO-US04929.

XX

PR
XX
PA
PA
PA
PA
PA
PA
PA
XX
PI
PI
XX
DR
XX
XX
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
XX
SQ

25-FEB-1999; 99US-0257417.
(GOUT/) GOUT I.
(RODN/) RODNIN N.
(FILO/) FILOENKO V.
(MATS/) MATSUKA G.
(SCAN/) SCANLAN M.
(OLDL/) OLD L.
(BILY/) BILYNSKY B.
Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
Bilynsky B;
WPI: 2000-572092/53.
Novel isolated nucleic acid molecules for diagnosing and treating melanoma, thyroid tumors, rectal, lung, breast and colon cancers -
Example 2; Page 62-63; 94pp; English.
The specification describes polynucleotides which are associated with melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5, Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for diagnosing and treating a patient with melanoma, thyroid tumour, rectal cancer, lung cancer, breast cancer or colon cancer. The present sequence represents a partial sequence of a Ki-67 nuclear antigen gene, and is a polynucleotide of the invention.
Sequence 861 BP; 271 A; 184 C; 205 G; 185 T; 16 other;

Query Match 3.5%; Score 344; DB 21; Length 861;
Best Local Similarity 99.6%; Pred. No. 1.4e-150;
Matches 444; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6490 gaaaaacagaaactgacccagcagcaagtgttaactggtagcaaggagcccaagaact 6549
|||||
Db 416 gaaaaacagaaactgacccagcagcaagtgttaactggtagcaaggagcccaagaact 475
QY 6550 cctaaaggaaaagcccaacccttagaagacttgctggctgaaagagctctccagaca 6609
|||||
Db 476 cctaaaggaaaagcccaacccttagaagacttgctggctgaaagagctctccagaca 535
QY 6610 ccagtagcactgacagccacgactcacgagaaaactaccacaaatagcctgcagatct 6669
|||||
Db 536 ccaatagcactgacagccacgactcacgagaaaactaccacaaatagcctgcagatct 595
QY 6670 ccaacaccgacccagtgggtaccacaaatcttcaagccacagtcacagagaagtctc 6729
|||||
Db 596 ccaacaccgacccagtgggtaccacaaatcttcaagccacagtcacagagaagtctc 655
QY 6730 aggaagcagcagtagagaaagaatccttagcactcagagaaacgaaacacatcagtagg 6789
|||||
Db 656 aggaagcagcagtagagaaagaatccttagcactcagagaaacgaaacacatcagtagg 715
QY 6790 aaagctatggacacaccccaaccagcagagggtgatgagaaacacatgaaagcatttatg 6849
|||||
Db 716 aaagctatggacacaccccaaccagcagagggtgatgagaaacacatgaaagcatttatg 775
QY 6850 gaaactccagtcagaaaattggacctgcaggaaaatttacctggcgcaaaagatggcca 6909
|||||
Db 776 ggaactccagtcagaaaattggacctgcaggaaaatttacctggcgcaaaagatggcca 835
QY 6910 caaaactcctaagaaaagcccaggc 6935
|||||
Db 836 caaaactcctaagaaaagcccaggc 861

RESULT 9

AAA64644
ID AAA64644 standard; DNA; 950 BP.

XX
AC AAA64644;

CC for use in e.g. forensics, paternity testing or for phenotypic typing for
CC diseases such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan
CC syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
CC familial hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
XX
SQ Sequence 201 BP; 35 A; 40 C; 48 G; 77 T; 1 other;

Query Match 1.4%; Score 134; DB 19; Length 201;
Best Local Similarity 99.5%; Pred. No. 3.7e-52;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2430 cttcagtcgacgaagtgcagcaaaacagccatctgataatgctctgcaagccctccctt 2489
DB 201 CTTCACTGTCACAGAAATGCAGCAAAACAGCCATCTGATAATGCTCTGCAAGCCCTCCCTT 142

QY 2490 aagacgagcagtgattagagaaaatgaaacgtagcaaaacgcccaggaacacacacaa 2549
DB 141 AGACGGCAGCTGTATTAGAGAAAATGGAAACGTAGCAAAAACGCCGGAACACCTTACAA 82

QY 2550 aatgactctctggagacaaaaacttcagatactgagagagagccttcaaaaacagatc 2609
DB 81 AATGACTTCTSTGGAGACAAAACCTTCAGATACCTGAGACAGAGCCCTTCAAAAACAGTATC 22

QY 2610 cactg 2614
DB 21 CACTG 17

RESULT 12
AAAF67522
ID AAF67522 standard; cDNA; 158 BP.
XX
AC AAF67522;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 3284.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHTR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI: 2001-091805/10.
XX

CC for use in e.g. forensics, paternity testing or for phenotypic typing for
CC diseases such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan
CC syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
CC familial hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
XX
SQ Sequence 201 BP; 35 A; 40 C; 48 G; 77 T; 1 other;

Query Match 3.1%; Score 299; DB 23; Length 1094;
Best Local Similarity 99.7%; Pred. No. 1.7e-129;
Matches 349; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8678 tggacacagaagtgttaattggcagcagagacgaagcaccctaaggaaagccccc 8737
DB 266 tggacgacagaagtgttaattggcagcagcagcagcaccctaaggaaagccccc 325

QY 8738 aaccctgaaagacccctgcagcttccaaagagctctcacaacaccagccacactgagg 8797
DB 326 aaccctgaaagacccctgcagcttccaaagagctctcacaacaccagccacactgagg 385

QY 8798 aactggcaaatggctgctgtagctttacaaagcctccaaagcaaacacactgcacagt 8857
DB 386 aactggcaaatggctgctgtagctttacaaagcctccaaagcaaacacactgcacagt 445

QY 8858 gaaacaccttaaaatccagagaggttcttcgggcccctaaagtacacccgtgggag 8917
DB 446 gaaacaccttaaaatccagagaggttcttcgggcccctaaagtacacccgtgggag 505

QY 8918 acgtgtaagcaccagacccctgtataatcacaagcaaaagcaaacacttccctgcccc 8977
DB 506 acgtgtaagcaccagacccctgtataatcacaagcaaaagcaaacacttccctgcccc 565

QY 8978 cactgcccctcaagagggaggtggcaaatggaagcgtcacgggaacc 9027
DB 566 cactgcccctcaagagggaggtggcaaatggaagcgtcacgggaacc 615

RESULT 11
AAAX10717/C
ID AAX10717 standard; DNA; 201 BP.
XX
AC AAX10717;
XX
DT 30-MAR-1999 (first entry)
XX
DE Human biallelic polymorphic DNA fragment WI-1011.
XX
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
XX
OS Homo sapiens.
XX
PN WO9820165-A2.
XX
PD 14-MAY-1998.
XX
PF 05-NOV-1997; 97WO-US20313.
XX
PR 06-NOV-1996; 96US-0030455.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Hudson T, Lander ES, Wang D;
XX
DR WPI: 1998-286974/25.
XX
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
XX
PS Claim 1; Page 75; 310pp; English.
XX
CC AAX10269-X12937 are human DNA fragments which contain biallelic
CC polymorphic markers which have been isolated using the primers
CC represented in AAX09121-X10268. The base occupying the polymorphic site
CC is indicated by the appropriate IUPAC-IUB ambiguity code. These fragments
CC can be used in methods for determining polymorphic forms in an individual
CC

PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 1036; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 158 BP; 54 A; 40 C; 34 G; 27 T; 3 other;

Query Match 1.1%; Score 110; DB 22; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.5e-41;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4699 aacttaactgcagcaagacgagcgtctacaactcctaaggaagccagcgtctagaa 4758
Db 11 aacttaactgcagcaagacgagcgtctacaactcctaaggaagccagcgtctagaa 70

Qy 4759 gacctggctgttaagagctctccagacacaggtcacactgagga 4808
Db 71 gacctggctgttaagagctctccagacacaggtcacactgagga 120

RESULT 13
AAAT20857
ID AAAT20857 standard; cDNA to mRNA; 219 BP.
XX
AC AAAT20857;
XX
DT 10-JUL-1996 (first entry)
XX
DE Human gene signature HUMGS02107.
XX
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
PN W09514772-A1.
XX
PD 01-JUN-1995.
XX
PF 11-NOV-1994; 94WO-JP01916.
XX
PR 12-NOV-1993; 93JP-0355504.
XX
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
XX
PI Matsubara K, Okubo K;
XX
DR WPI; 1995-206931/27.
XX
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

XX
PS Claim 1; Page 743; 2245pp; Japanese.
XX
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
SQ Sequence 219 BP; 93 A; 28 C; 57 G; 40 T; 1 other;

Query Match 0.8%; Score 78; DB 16; Length 219;
Best Local Similarity 100.0%; Pred. No. 6.5e-26;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9659 agagagtaacgcgcgagtggtcaagaggtgtgcagaaaaatccaaagaggtgaggaatg 9718
Db 58 agagagtaacgcgcgagtggtcaagaggtgtgcagaaaaatccaaagaggtgaggaatg 117

Qy 9719 tgtgtgtcaagaaaaataa 9736
Db 118 tgtgtgtcaagaaaaataa 135

RESULT 14
AAS69163
ID AAS69163 standard; cDNA; 1126 BP.
XX
AC AAS69163;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4967.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG04976.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 4967; 103pp; English.
XX

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1126 BP; 371 A; 274 C; 287 G; 194 T; 0 other;

Query Match 0.7%; Score 64; DB 23; Length 1126;
Best Local Similarity 100.0%; Pred. No. 2.5e-19;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8752 ctggccagcttccaagagctctctcaaacaccagccacactgaggaaactggcaaatggt 8811
|||||

Db 379 ctggccagcttccaagagctctctcaaacaccagccacactgaggaaactggcaaatggt 438
|||||

QY 8812 gctg 8815
|||||

Db 439 gctg 442
|||||

RESULT 15

AAL27414

ID AAL27414 standard; DNA; 51 BP.

XX AAL27414;

DT 24-JAN-2002 (first entry)

XX Human SNP oligonucleotide #622.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.

XX Homo sapiens.

OS WO200147944-A2.

PN 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

XX ShImkets RA, Leach M;

XX WPI; 2001-465210/50.

XX

PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -

PS Claim 1; Page 1565; 4143pp; English.

XX

CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.

XX

SQ Sequence 51 BP; 20 A; 16 C; 8 G; 7 T; 0 other;

Query Match 0.5%; Score 51; DB 22; Length 51;

Best Local Similarity 100.0%; Pred. No. 2.9e-13;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4433 tatgcactgacagccacgactcacgagaaaactaccacaaatagctgcga 4483
|||||

Db 1 tatgcactgacagccacgactcacgagaaaactaccacaaatagctgcga 51
|||||

Search completed: August 24, 2002, 01:37:14

Job time: 28334 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2002, 17:45:00 ; Search time 9568.17 Seconds
(without alignments)
13776.016 Million cell updates/sec

Title: US-09-700-906A-1_COPY_197_9962
Perfect score: 9766
Sequence: 1 atgtggccacgagacgctt.....tcattagggacagtgaagata 9766

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_gss.*
- 13: em_gss_hum.*
- 14: em_gss_inv.*
- 15: em_gss_pln.*
- 16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	731	7.5	1058	10	BG389162 602414982
2	703	7.2	1084	10	BM455229 AGENCOURT
3	691	7.1	1022	10	BM474441 AGENCOURT
4	640	6.6	1009	10	BM459184 AGENCOURT
5	619	6.3	790	9	AL556595 AGENCOURT
6	576	5.9	1070	10	BM455791 AGENCOURT
7	575	5.9	697	10	BG403181 602418615
8	553	5.7	1068	10	BM452654 AGENCOURT
9	552	5.7	606	10	BE395002 601308294
10	544	5.6	580	9	AW835684 QV4-LT001
11	537	5.5	1047	10	BM476598 AGENCOURT
12	535	5.5	822	9	AU140843 AGENCOURT
13	525	5.4	753	10	BI868409 603392387
14	519	5.3	751	10	BG252689 602366335
15	510	5.2	692	10	BG677474 602625107
16	500	5.1	558	9	AW503684 UI-HF-BN0
17	497	5.1	546	9	AW502198 UI-HF-BN0

18	491	5.0	758	10	BG393561	BG393561 602411954
19	487	5.0	824	10	BE540030	BE540030 601060802
c 20	481	4.9	596	10	BF516576	BF516576 UI-H-BW1-
21	475	4.9	584	12	AQ471476	AQ471476 CITBI-E1-
22	471	4.8	471	10	BF898662	BF898662 QV1-MT022
23	459	4.7	596	9	AW502251	AW502251 UI-HF-BN0
24	442	4.5	852	10	BG613631	BG613631 602641986
c 25	439	4.5	599	9	A1064694	A1064694 HA0443 Hu
c 26	435	4.5	436	9	AA262266	AA262266 zs25e07.s
27	432	4.4	907	10	BG164890	BG164890 602343508
28	430	4.4	820	9	AU130074	AU130074 AU130074
29	421	4.3	505	9	AW504977	AW504977 UI-HF-BN0
30	421	4.3	683	9	AW853584	AW853584 RCL-CT025
c 31	420	4.3	557	9	AW835681	AW835681 QV4-LT001
c 32	416	4.3	446	9	A1494288	A1494288 QY98D08.X
33	414	4.2	561	9	AW499743	AW499743 UI-HF-BR0
34	414	4.2	667	10	BI861849	BI861849 603388877
35	413	4.2	639	9	BE083942	BE083942 PMO-BT065
36	411	4.2	672	9	AW853570	AW853570 RCL-CT025
37	406	4.2	471	9	AA262877	AA262877 zs25e07.r
38	403	4.1	851	10	BF966006	BF966006 602277193
c 39	401	4.1	465	9	AA809977	AA809977 oa95q07.s
40	397	4.1	960	10	BG032883	BG032883 602300466
41	396	4.1	508	9	A1002455	A1002455 oq82h11.s
42	394	4.0	688	10	BE740099	BE740099 601595030
c 43	388	4.0	912	9	AL562952	AL562952 AL562952
44	387	4.0	582	9	BE084144	BE084144 PMO-BT065
45	383	3.9	440	9	AW501940	AW501940 UI-HF-BR0

ALIGNMENTS

RESULT 1

BG389162 602414982F1 NIH_MGC_92 1058 bp mRNA linear EST 12-MAR-2001
LOCUS 602414982F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4523108 5',
DEFINITION mRNA sequence.

ACCESSION BG389162

VERSION BG389162.1 GI:13282608

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1058)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10425 row: d column: 21

High quality sequence stop: 752.

Location/Qualifiers

1..1058

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4523108"

/tissue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;

Site: 2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT	404 a	258 c	252 g	144 t
Query Match	7.5%; Score 731; DB 10; Length 1058;			
Best_Local_Similarity	100.0%; Pred. No. 0;			
Matches 731; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
Qy 2653	cagaagctcacctgttgaaagtgaagagtgaaagacacaaatacagaataatgttgagtcac 2712			
Db 1	CAGAAGCTTACCTGTGTGAAAGTAAGAGTGAAGAAACAATAACAGAATAATTGTTGAGTGCATC 60			
Qy 2713	ctaaaaagaggttcagaaagcgcaactactacaacaaagagagaagagatgaagaa 2772			
Db 61	CTAAAAAGAGGTTCAGAAGGCAACACTACTACAACAAGAGAGAGAGATGAAGGAA 120			
Qy 2773	atagaaagacacttttgagacataaagaaataattgattaaaaagaaacagatgaaag 2832			
Db 121	ATAGAAAGACCTTTTGAAGACATATAAGGAAATAATTGAAATTAAGAAACGATGAAAG 180			
Qy 2833	atgaagcaatgaagagatcaagaacttgggggcagaatgtgcaccactgtctgacctg 2892			
Db 181	ATGAAGCAATGAGAGATCAAGAACTTGGGGCAGAAATGTGCACCAATGTCTGACCTG 240			
Qy 2893	acagacctcaagagcttgcctgatacagaaactcatgaagacacgacgctggccagaat 2952			
Db 241	ACAGACCTCAAGAGCTTGCTTCATACAACTCATGAAAGACACGGCAGCTGGCCAGAAT 300			
Qy 2953	ctctccaaacccaaagatcatgcgaagcgacaaagagtgagaaaggcaaatcaactaaa 3012			
Db 301	CTCTCTCCAAACCCCAAGATCATGCCAAGGCACCAAGAGTGAGAAAGGCCAAATCACTAAA 360			
Qy 3013	atgcctcgcagtcattacaaccagaacccaataaacacccccacacacacacacacag 3072			
Db 361	ATGCCCTGCCAGTCATTACACACCAACCAATAAACACCCCAACACACACAAACACAG 420			
Qy 3073	ttgaagcatccctggggaaagtgtgtgaaagaagagctcctagcagtcggcaagttc 3132			
Db 421	TTGAAGGCATCCCTGGGGAAAGTAGGTGTGAAAGAAGAGCTCCTAGCAGTCGGCAAGTTC 480			
Qy 3133	acaggaagcttcagggagacacgacacagcagcagagagaccagcagagatggcaagagc 3192			
Db 481	ACACGGACGTCAGGGGAGACACCCACACGACAGAGAGCCAGCAGGAGATGGCAAGAGC 540			
Qy 3193	atcagaagctttaaggagctctccaaagcagatccttgaccacgacgcccgtgtaactgga 3252			
Db 541	ATCAGACGTTTAAAGAGTCTCCAAGCAGATCTCGACCCACGACGCCGCTGTAACTGGA 600			
Qy 3253	atgaagaagtgcccaagaaacgcctaaagaaagagcccgactcactagaagacctggctgac 3312			
Db 601	ATGAAGAAGTGGCCAAAGAACGCTTAAGGAAGAGGCCAGTCACTAGAAGACCTGGCTGGC 660			
Qy 3313	ttcaagaagctcttcagaaacccaggtccctctcaggaatcaatgactgatgaaaaaact 3372			
Db 661	TTCAAAAGAGCTTTCACAGACACAGGTCCTCTGAGGAATCAATGACTGTATGAGAAAACT 720			
Qy 3373	accataatagc 3383			
Db 721	ACCAAAATAGC 731			
RESULT 2				
BM455229				
LOCUS	BM455229			
DEFINITION	AGENCOURT_6405486 NIH_MGC_85 Homo sapiens cdna clone IMAGE:5500179			
ACCESSION	BM455229			
VERSION	BM455229.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			

[illegible]

```
Qy 8777 aaacaccaggccacactgaggaaactggcaaatgggtgctgctgatatagctttcaaacgctc 8836
|||||
Db 361 AAACACCAGGCGCACACTGAGGAACCTGGCAAAATGGTGTCTGTATAGCTTTTACAAGCGCTC 420

Qy 8837 caaagcaaacctgacagtggaaaaacctctaaaaataatcagaagaagtcttcggggccc 8896
|||||
Db 421 CAAAGCAAACCTGACAGTGGAAACCTCTAAAAAATATTCAGAAAGAGTTCTTCGGGGCC 480

Qy 8897 ctaagtagaacccgtgggagactgtggaagcaccagagagaccctgtaaaaatcacaaagca 8956
|||||
Db 481 CTAAGTAGAACCCGTGGGAGACGTGGTAAGCACAGAGACCCCTGTAAATACAAAGCA 540

Qy 8957 aaagcaacacttcctgccccactgccccttcaagagaggggggtggcgaagatggaagcg 9016
|||||
Db 541 AAAGCAACACTTCCCTGCCCTCCACTGCCCTTCAAGAGGGAGGTGGCAAGATGGAAAGCG 600

Qy 9017 tcaggggaaccaagaggtgcgtgcatgccagcaccagagagaaatttgaggagctgc 9076
|||||
Db 601 TCAGGGGAACCAAGAGGTGGCGTGCATGCCAGCACCAAGAGGAATTTGGAGGAGCTGC 660

Qy 9077 cagccagcaagaagcagaggggtgtctccaggggcaagaggcaaatcatccgaacccgtgg 9136
|||||
Db 661 CAGCCAGCAAGACAGAGGCTTGTCTCCAGGGCAAGAGGCAAAATCATCCGAACCCGTGG 720

Qy 9137 tcataatgaagagaagttgaggactctgcacaaagaattgaacctgcgga 9188
|||||
Db 721 TCATCATGAAGAAGTTTGAGGACTTCTGCAAAAAAAGAAATTGAACCTGCGGA 772

RESULT 6
BM455791
LOCUS BM455791 1070 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT 6412973 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5498229
5', mRNA sequence.
ACCESSION BM455791
VERSION BM455791.1 GI:18504831
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1070)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaops-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM2129 row: j column: 22
High quality sequence stop: 585.
Location/Qualifiers
1..1070
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5498229"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 386 a 291 c 238 g 155 t
ORIGIN
```

```
Query Match 5.9%; Score 576; DB 10; Length 1070;
Best Local Similarity 99.6%; Pred. No. 1.7e-259;
Matches 726; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5063 caagtctaaactggcagcagaagcggcgtgagaactcctaagggaaagtctgaagtcctc 5122
|||||
Db 1 CAAGTCTTAACCTGGCAGCAAGAGGCAGCTGAGAACTCCTAAGGGAAAGTCTGAAGTCCCTCG 60

Qy 5123 aagacctggccggtctcatcgagctcttcagacaccacaaagtcacactcaaggaataatga 5182
|||||
Db 61 AAGACTCGCCCGCTTCATCGAGCTTCCACACACCAAGTCACACTAAGGAATCAATGA 120

Qy 5183 ctaatgaaaaaactaccaaaagtatctactacagagcttcacagccagacctagtggacacc 5242
|||||
Db 121 CTAACGAAAAAACTACCAAAGTATCCTACAGAGCTTCACAGCCAGACCTAGTGGACACC 180

Qy 5243 caacaagctccaagccacagcccaagagaagtcctcaggaagaacagacacactgaagaaga 5302
|||||
Db 181 CAACAAGCTCCAAGCCACAGCCCAAGAGAGTCTTCAGGAAGCAGACACTGAAGAAGAAT 240

Qy 5303 ttttagcttaggaacaaacgccatcagcagggcaaaagccatgcacacacacccaaaccag 5362
|||||
Db 241 TTTTAGCATTTAGGAAACAAACGCCATCAGCAGGCAAAAGCCATGCACACACCCAAACAC 300

Qy 5363 cagtaggtgaagaaaaagacatcaacacgctttttgggaactccagtcgagaacttgacc 5422
|||||
Db 301 CAGTAGGTGAAGAGAAAGACATCAACACGTTTGGGAACCTCCAGTGCAGAACTGGACC 360

Qy 5423 agccaggaaatttaactggcagcaatagacggtctacaaactcgtgaaggaaagccaggg 5482
|||||
Db 361 AGCCAGGAAATTTAGCTGGCAGCAATAGACGGCTACAAACTCGTAGGAAAAAGGCCAGG 420

Qy 5483 cctcagaagaactgactggcttcagagactttccagacacatgcactgataacccca 5542
|||||
Db 421 CTCTAGAAGAACTGACTGGCTTCAGAGAGCTTTTCCAGACACCATGCATGATAACCCCA 480

Qy 5543 cagctgatgagaaaaactaccaaaaataactctgcaaatctccgcaatcagacccagcg 5602
|||||
Db 481 CGACTGATGAGAAAACTACCAAAAAAATACTCTGCAAAATCTCGCAATCAGACCCAGCG 540

Qy 5603 acaccccaacaaacaaacaaagcagcccaagagaagcctcaagaagcagacgtagagg 5662
|||||
Db 541 ACACCCCAACAAACAAACAAAGCAACGCCCAAGAGAGCCTCAAGAAAGCAGACGTAGAGG 600

Qy 5663 aagaatttttagcattcaggaactaaacaccatcagcagcagcaagccatgcacagccta 5722
|||||
Db 601 AAGAATTTTACCATTCAGAAACTAACACCATCAGCAGGCAAAAGCCATGCACACCCCTA 660

Qy 5723 aagcagcagtagtgaagagaaagacatcaacacatttgtgggactccagtggaagaaac 5782
|||||
Db 661 AAGCAGCAGTAGTGAAGAGAAAGACATCAACACATTTGTGGGGACTCCAGTGGAGAAAC 720

Qy 5783 tgaacctgc 5791
|||||
Db 721 TGGACCTGC 729

RESULT 7
BG403181
LOCUS BG403181 697 bp mRNA linear EST 12-MAR-2001
DEFINITION 602418615F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4525994 5',
mRNA sequence.
ACCESSION BG403181
VERSION BG403181.1 GI:13296629
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 697)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
```


Qy	5906	aatcaatgaccgatgacaaatcagagaagtatcctctgcaaatctccaaaccagaccag	5965
Db	405	AATCAATGACCGATGACAAATACAGAAGTATCCTTGC	464
Qy	5966	tcaaaaccccaacaagctccaagcaacgactcaagatatccttggggaagtagtggtga	6025
Db	465	TCAAAACCCACACAAAGCTCCAAGCAGACTCAAGATATCCTTGGGGAAAGTAGTGTGA	524
Qy	6026	aagaagaggtctaccagtcgggcaagctcacacagacgtcaggaagaccacacagacac	6085
Db	525	AAGAAGAGGTCTTACCAGTCGGCAAGCTCACACAGAGCTCAGGGAAGACACACAGACAC	584
Qy	6086	acagagagacagcaggagatggaagagatcaaaagcgcttaaggaatctgcgaaagcaga	6145
Db	585	ACAGAGACAGCAGGAGATGGGAAGAGACTCAAGAGCGTTTAAAGGAATCTGCAAGCAGA	644
Qy	6146	tgctggaccacagaaactatggaaactggatgagagtggtgccaagacaacacctaa	6200
Db	645	TGCTGGACCCAGCAAACTATGGAACCTGGATGGAGAGTGCCCAAGAACAACCTAA	699
RESULT 9			
BE395002	LOCUS BE395002 606 bp mRNA linear EST 21-JUL-2000		
DEFINITION	601308294F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3626638 5', mRNA sequence.		
ACCESSION	BE395002		
VERSION	BE395002.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLCM308 row: k column: 23 High quality sequence stop: 602.		
FEATURES	Location/Qualifiers		
source	1. 606 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3626638" /clone_lib="NIH_MGC_44" /tissue_type="endometrium, adenocarcinoma cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: uterus; Vector: pOPB7; Site:1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."		
BASE COUNT	231 a 158 C 144 G 73 T		
ORIGIN			
Query Match	5.7%; Score 552; DB 10; Length 606;		
Best Local Similarity	99.8%;		
Matches	602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
Qy	7901	aagaaccagcagcggtgatgaggcgatcaaatgattgaagcaactggtcaagaagaac	7960

Db	1	AAGAACCCAGCAAGCGGTGATGAGGGCATCAAAAGTATTGAAGCAACAGTGCAAAGAAAGAAAC	60
Qy	7961	caaacccagtagaagagaaacccagcagaggaagccaaagagcagcctaaagaaagcccc	8020
Db	61	CAAAACCCAGTAGAGAGGAACCCAGCAGGAGGAAGCCCAAGACACCTTAAAGAAAGAGGCC	120
Qy	8021	aacccctggaaagcctggccggcttcacagagctctctgaaacatcaggtccacactcagg	8080
Db	121	AACCCCTGGAAAGACCTGGCCGGCTTCACAGAGCTCTCTGAAACATCAAGTCCACACTCAGG	180
Qy	8081	aatcactgactgtgcgaagccactaaaataacctcggaatctccccactagaagtgg	8140
Db	181	AATCACTGACTGCTGGCAAGAACCACTAAATACCTCGGAATCTCCCCCACTGAAAGTGG	240
Qy	8141	tagacaccacagcagcacaaagaggtcatctcagcacacgtgtgcagaaggtacaagtaa	8200
Db	241	TAGACACACAGCAAGCACAAAGAGGCATCTCAGGACACAGTGTGCAGAAAGTACAAGTAA	300
Qy	8201	aagaagagccttcagcagctcaagttcacacacaaacatcagggggaaacccagcgatcgacac	8260
Db	301	AAGAAGAGCCTTCAGCAGTCAAGTTCACACAAACATCAGGGGAAACCCAGCATGCAGACA	360
Qy	8261	aagaaccagcaggtggaagataaaggatcaaacagcatgaaagaaatctgcaaacagacac	8320
Db	361	AAGAACCAGCAGGTGAAGTAAAGGCATCAAGCATTTGAAGGAATCTGCAAAACAGACAC	420
Qy	8321	cggtccagcagcagtgtaactggcagcagagagacggccaagagcagccaggggaaagtg	8380
Db	421	CGGCTCAGCAGCAAGTCTAAGTGGCAGCAGGACACAGCCAAAGACACCCAGGAAAGTG	480
Qy	8381	cccagccatagaagacctagctggcttcaagaccagcagcaggtgcacactgaaagt	8440
Db	481	CCCAAGCCATAGAAGACCTAGCTGGCTTCAAGACCCAGCAGCAGGTGCACACTGAAGAA	540
Qy	8441	caatgactgatgacaaacaccactaaaataacctgcaaatcatcaccagaactagaagaca	8500
Db	541	CAATGACTGATGACAAACCCACTAAATACCTGCAATATCATCAGCAAGTGAAGACA	600
Qy	8501	ccg 8503	
Db	601	CCG 603	
RESULT	10		
LOCUS	AW835684/c	580 bp	linear
DEFINITION	QV4-LT0016-140200-105-h06	LT0016	Homo sapiens
ACCESSION	AW835684		
VERSION	AW835684.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 580)		
	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,		
	Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,		
	Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,		
	Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare		
	,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and		
	Simpson,A.J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed		
	sequence tags		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (7),	3491-3496 (2000)
MEDLINE	20202663		
COMMENT	Contact: Simpson A.J.G.		
	Laboratory of Cancer Genetics		
	Ludwig Institute for Cancer Research		
	Rua Prof. Antonio Prudente 109, 4 andar,		01509-010, Sao Paulo-SP,
	Brazil		
	Tel: +55-11-2704922		
	Fax: +55-11-2707001		
	Email: asimpson@ludwig.org.br		


```
|||||
Db 421 GAAATTTTAGCATTCAGAAATTAACACCATCAGCAGCAAGCCATCACACGCTTAA 480
Qy 5725 gcagcagtagtggaagaaagacatacaacatttggggactccagtgagaaactg 5784
Db 481 GCAGCAGTAGTGAGAGAAAGACATCAACACATTTGTGGGGCTCCAGTGGAGAACTG 540
Qy 5785 gacctgctaggaatttacctggcagcaagagagcgccacaaaactcctaaagaaagccc 5844
Db 541 GACCTGCTAGGAAATTTACCTGGCAGCAAGAGACGGCCACAAACTCCTTAAAGAAAGGCC 600
Qy 5845 aagcctctagagatctgctggcttcaagagagctctccagacaacacaggtcaactgag 5904
Db 601 AAGGCTCTAGAGAGATCTGGCTGGCTTCAAGAGGCTCTTCCAGACACACAGGTCACACTGAG 660
Qy 5905 gaatcaatgaccgatgacaaatcacagaagtatcctgcgaatctccacaacccagacc 5964
Db 661 GAATCAATGACCGATGACCAATACAGAGATATCTTGCAAAATCTCCACAACGACCCA 720
Qy 5965 gtcaaaaccccaacaagctcc 5985
Db 721 GTCAAAACCCCAACAGCTCC 741

RESULT 12
AUI40843
LOCUS AUI40843 PLACE4 Homo sapiens cDNA clone PLACE4000347 5', mRNA
DEFINITION AUI40843
ACCESSION AUI40843
VERSION AUI40843.1 GI:11002364
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 822)
Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isogai,T.)
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE4000347"
/clone_lib="PLACE4"
/tissue_type="placenta"
/note="vector: pME18SFL3"
BASE COUNT 305 a 211 c 177 g 125 t 4 others
ORIGIN
Query Match 5.5%; Score 535; DB 9; Length 822;
Best Local Similarity 99.6%; Pred. No. 3,2e-240;
Matches 685; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 4941 agcagttggcagctcacagacatcaggagagactacacacacacagagcccaac 5000
Db 1 AGCAGTTGGCAAGCTCACAGACATCAGGAGAGACTACACACACACAGAGCCCAAC 60
```

```
Qy 5001 aggaatggttaagacatgaaagcatttatgagtgctccaaagcagatcttagactcaagc 5060
Db 61 AGGAGATGGTAAGACGATGAAGCATTTATGGAGTCTCCAAAGCAGATCTTAGACTCAGC 120
Qy 5061 agcaatgcttaactggcagcaagagcagctgagaactcctaaaggaaagtcgtgaagtccc 5120
Db 121 AGCAAGTCTAACTGGCAGCAAGAGCGCAGCTGAGAACTCTCTAAGGGAAAGTCTGAAAGTCCC 180
Qy 5121 tgaagacctgcccgttcacatcgagctctccagacacacaagctcaactaagaatacaat 5180
Db 181 TGAAGACCTGGCCGGCTTCATCGAGCTCTTCAGACACACCAAGTCCACTAAGAGATTCAAT 240
Qy 5181 gactaatgaaaaaactcacaagatcctacagagcttcacagccagacacctagtggacac 5240
Db 241 GACTAACGAAAAAACTACCAAGATGTTCTACAGAGCTTCACAGCCAGACCTAGTGGACAC 300
Qy 5241 cccaacagctccagccacagcccaagagaagtctcaggaaagcagacactgaagaaga 5300
Db 301 CCCAACAAAGCTCCAAGCCACAGCCCAAGAGAAAGTCTCAGGAAAGCAGACACTGAAGAAGA 360
Qy 5301 atttttagcatttagaaacaaacccatcagcagcagcaagccatgcacacaccccaaac 5360
Db 361 ATTTTTCATTTTGGAAACAAACGCCCATCAGCAGGCAAGCCATGCACACACCCCAAGCC 420
Qy 5361 agcagtagtgtaagagaagacatcaacacgctttttgggaactccagtgcaagaactgga 5420
Db 421 AGCAGTAGTGAAGAGAAAGACATCAACACGTTTGTGGAACTCCAGTGCAGAACTGGA 480
Qy 5421 ccagccaggaatttacctggcagcaatagacggtcacaactcgttaaggaaagccca 5480
Db 481 CCAGCAGGAAATTTACCTGGCAGCAATAGACGGGTACAAACTCGTAAAGGAAAGGCCCA 540
Qy 5481 ggcttaagaaactgactggcttcagagagctttccagacacaccatgcactgataaac 5540
Db 541 GGCTCTAGAAGAACTTGACTGGCTTCAGAGAGCTTTTCCAGACACCATGCACGTGATAACCC 600
Qy 5541 cacagctgatgagaaactaccaaaaatactctgcaaatctccgcaatccgcaatccagccagc 5600
Db 601 CACAACCTGATGAGAAACTACCAAAAAATACTCTGCAAAATCTCCGCAATCAGACCCAGC 660
Qy 5601 ggcaccccccaacacacaaagcaacgg 5628
Db 661 GGACACCCCAACAAACACAAAGCAACGG 688

RESULT 13
BI868409
LOCUS 603392387F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402684 5',
DEFINITION mRNA sequence.
ACCESSION BI868409
VERSION BI868409.1 GI:16042082
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 753)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12026 row: m column: 21
High quality sequence stop: 741.
```

FEATURES		Location/Qualifiers	
source	1..753	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5402684" /clone_lib="NIH_MGC_90" /tissue_type="adenocarcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	
BASE COUNT	286 a	144 c	192 g 131 t
ORIGIN			
Query Match 5.4%; Score 525; DB 10; Length 753; Best Local Similarity 99.8%; Pred. No. 1.6e-235; Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	9161	ctctcgaaagaattgaacctgcggaagagctgaacagcaacgacatgaataccaa	9220
Db	13	CTTCTCGAAAAAGAAATTGAACCTGCGAGAGAGCTGAACAGCACAGACATGAAACCAACA	72
QY	9221	aagaggaaacacaaattacaaagactcgtccctgaaataaagggaatatccctgcgctcca	9280
Db	73	AAGAGGAACACAAATTACAAAGACTCGTCCCTCGAAATAAGGGAATATCCCTGCGCTCCA	132
QY	9281	gagcccaagataagactgaggagagacagcaataactaggtctttgtattagcagaaa	9340
Db	133	GAGGCCAAAATAAGACTGAGGCGAAGACAGCAATAACTCAGGTCTTTGTATTAGCAGAAA	192
QY	9341	gaatagaatacaacagaaatgaagaagcccatgaagacctcccagagatggacattc	9400
Db	193	GAATAGAAATAACAAATGAAAGAGCCCATGAAGACCTCCCCAGAGATGGACATTC	252
QY	9401	agaatccagatgatgagcccggaacccataccttagagacaaaagtcaactgagaacaaa	9460
Db	253	AGAATCCAGATGATGAGCCCGGAAACCCATACCTAGAGACAAAGTCACTGAGAACAAAA	312
QY	9461	ggtcgttaggtctgttagacagaataagagctcccagcctaaagtgaggagagagcg	9520
Db	313	GGTGTGTGAGGTCTGTAGACAGAAATGAGAGTCTCCAGGCTTAAGGTGGCAGAGGAGCG	372
QY	9521	gagggcagaagagtgcgaagttctcatgcagaatcagaaagggaagagcaggaa	9580
Db	373	GAGGGCAGAAGAGTGCAGAGTTCTCATGCAGAAATCAGAAAGGGAAGAGAGACAGGAA	432
QY	9581	attcagactccatgtgcctgagatcaagaaagacaaaaagccgctgcagcaagcactt	9640
Db	433	ATTTCAGACTCCATGTGCCTGAGATCAAGAAAGACAAAAAGCCAGCCTGCAGCAAGCACTT	492
QY	9641	tggagagcaaatctgtgacagagagtaacgcggagtctcaagaggtgtgcagaaaaatccaa	9700
Db	493	TGGAGAGCAAAATCTGTGCAGAGAGTAAACGCGAGTGTCAAGAGGTGTGCAGAAAAATCCAA	552
QY	9701	agagagctgaggacaatgtgtgtgtcaagaaaaataa	9736
Db	553	AGAAGGCTGAGGACAAATGTGTGTCAAGAAATAA	588
RESULT 14			
BG252689			
LOCUS			
DEFINITION			
602366335F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4474376 5', mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 751)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10298 row: f column: 09
High quality sequence stop: 667.

FEATURES
source

1..751
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4474376"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
273 a 198 c 171 g 109 t

BASE COUNT
ORIGIN

Query Match 5.3%; Score 519; DB 10; Length 751;
Best Local Similarity 99.7%; Pred. No. 1.1e-232;
Matches 619; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	6642	gaaactacacaaatagcctgcagatctccacacacagaccctagtggtaccaccaaat	6701
Db	1	GAAACTACCAAAATAGCCTGCAGATCTCCACAACAGACCCAGTGGGTACCCCAACAAT	60
QY	6702	cttcagagccacagtcacagagaagtcctcaggaaagcagacgtagagaagaatccctagc	6761
Db	61	CTTCAAGCCACAGTCTCCAGAGAGTCTCAGGAAAGCAGATGTAGAGGAAAGAAATCCTTAGC	120
QY	6762	actcagaaacgaacaccatcagtaggaagctatggacacaccccaaacccagcagagg	6821
Db	121	ACTCAGAAACGNACACCATCAGTAGGGAAAGCTATGGACACACACCCAAACCCAGCAGG	180
QY	6822	tgatgagaagacatgaaagcatttatgggaactccagtcagaaattggacctgccag	6881
Db	181	TGATGAGAAAGACATCAAGCATTATGGGAACCTCCAGTGCAGAAATTTGACCTGCCAGG	240
QY	6882	aaatttacctggcagcaaaagatggccacaaactcctaaggaagaaagccagcctctaga	6941
Db	241	AAATTTTACCTGGCAGCAAAAGATGGCCCAAACTCTTAAGGAAAGAGCCAGGCTCTAGA	300
QY	6942	agacctggctggcttcaagagcctctccagacacacagcagcactgacaaagccacgactga	7001
Db	301	AGACCTGGCTGGCTTCAAGAGCTCTTCCAGACACCCAGGCACACTGCACAGCCACGACTGA	360
QY	7002	tgagaaactacaaaatagcctgcaaatctccacacacagaccctagtgagacccccagc	7061
Db	361	TGAGAAAACTACCAAAATAGCCTGCAAAATCTCCACAACAGACCCAGTGCACACCCAGC	420
QY	7062	aagcacaagacacggcccaagaaacctcaggagaagcagacgtagagagaatctttt	7121
Db	421	AAGCACAAAGCAACGCCCCCAAGAGAAACCTCAGGAAAGCAGACGTTAGAGGAATTTT	480
QY	7122	agcactcaggaagaaacacccatcagcagggcaagcattggacacaccccccaaacccagcag	7181
Db	481	AGCACTCAGAAACGAAACCATCAGCAGGCAAGGCGATGGACACACACCAAAACCCAGCT	540

QY 7182 aagtatgagaaaaatatacaacacatttgtggaactccagtgcagaaactggacctgct 7241
|||||
Db 541 AAGTGATGAGAAAAATATCAACACATTTTGTGGAACCTCCAGTGCAGAAACTGGACCTGCT 600
QY 7242 aggaatttacctgcagcaa 7262
|||||
Db 601 AGGAAATTACCTGGCAGCAA 621

RESULT 15
BG677474

LOCUS BG677474 692 bp mRNA linear EST 01-MAY-2001
DEFINITION 602625107f1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4750249 5',
mRNA sequence.
ACCESSION BG677474
VERSION BG677474.1 GI:13908871
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1A10604 row: m column: 02
High quality sequence stop: 690.
Location/Qualifiers
1..692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4750249"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 247 a 170 c 174 g 101 t
ORIGIN

Query Match 5.2%; Score 510; DB 10; Length 692;
Best Local Similarity 100.0%; Pred. No. 1.8e-228; Mismatches 0; Indels 0; Gaps 0;
Matches 510; Conservative 0;

QY 8752 ctggccagcttccaagagctctctcaaacaccagccacactgaggaactggcaaatggt 8811
|||||
Db 40 CTGGCCAGCTTCCAAGAGCTCTCTCAACACCAGGCCACACTGAGGAACTGGCAATGGT 99

QY 8812 gctgtgtagctttacaagcgctccaagaacacacacctgacagtggaaaacctctaaaa 8871
|||||
Db 100 GCTGTGTAGCTTTACAAGCGCTCTCAAGCAACACCTGACAGTGGAAACCTCTAAAA 159

QY 8872 atatccagaaagtcttcggggccctaaagttagaccctgggagacgtggttaagacc 8931
|||||
Db 160 ATATCCAGAAGAGTTCCTTCGGGGCCCTTAAGTAGAACCCGTGGGAGACGTGGTAAGCACC 219

QY 8932 agagacctgtataatcaagaagcaaacacttccctgccccccactgccccttcag 8991
|||||
Db 220 AGAGACCTGTAAATACAAAGCAAAAGCAACACTTCCCTGCCCCACTGCCCTTCAAG 279

QY 8992 aggggaggtggaagcgctcacgggaaccaagaggtcgctgcatgccagca 9051

Db 280 AGGGGAGGTGGCAAAAGATGGAAGCGCTACGGGAACCAAGAGGCTGCGCTGCATGCCAGCA 339
QY 9052 ccagaggaatttgtgagagagctgccagccagcaagaagcadaggtttgtctccagggca 9111
|||||
Db 340 CCAGAGGAAATTGTGAGGAGCTGCCAGCCAGCAAGAGAGGTTGCTCCAGGGCA 399
QY 9112 agaggcaaatcatccgaaccccggtggtcatcatgaagaagaagtgttggaggactcttcgcaaa 9171
|||||
Db 400 AGAGGCAAAATCATCCGAACCCGCTGGTTCATCATGAAGAGAAAGTTTGAGGACTTCTGCAAAA 459
QY 9172 agaattgaacctgcgggaagagctgaacagcaacgacatgaaacccaacaaagagggaacac 9231
|||||
Db 460 AGAATTGAACCTGCCGGAAGAGCTGACAGCAACGACATGAAAAACCAACAAAGAGGAACAC 519
QY 9232 aaatlacaagactcgggtccctgaaaaataag 9261
|||||
Db 520 AATTACAAGACTCGGTCCCTGAAAAATAAG 549

Search completed: August 23, 2002, 20:25:27
Job time: 9627 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2002, 17:45:00 ; Search time 251.89 Seconds
(without alignments)
9523.430 Million cell updates/sec

Title: US-09-700-906A-1_COPY_197_9962
Perfect score: 9766
Sequence: 1 atgtggccacgagcgcct.....tcataggacagtgaagata 9766

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/1na/5A-COMB.seq.*
2: /cgn2_6/ptodata/2/1na/5B-COMB.seq.*
3: /cgn2_6/ptodata/2/1na/6A-COMB.seq.*
4: /cgn2_6/ptodata/2/1na/6B-COMB.seq.*
5: /cgn2_6/ptodata/2/1na/PCTUS-COMB.seq.*
6: /cgn2_6/ptodata/2/1na/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

COMPARISONS

Result No.	Score	Query Match	Length	ID	Description
1	389	4.0	561	4 US-09-328-111-458	Sequence 458, Appli
C 2	20	0.2	340	1 US-08-599-252-105	Sequence 105, Appli
C 3	20	0.2	340	5 PCT-US96-06352-105	Sequence 105, Appli
C 4	20	0.2	340	5 PCT-US96-06583-105	Sequence 105, Appli
C 5	20	0.2	530	4 US-08-991-789A-221	Sequence 221, Appli
C 6	20	0.2	530	4 US-09-062-451-221	Sequence 221, Appli
7	20	0.2	7881	2 US-08-751-189-1	Sequence 1, Appli
8	20	0.2	7881	2 US-09-060-836-1	Sequence 1, Appli
9	20	0.2	7881	4 US-09-184-445-1	Sequence 1, Appli
10	19	0.2	259	1 US-08-222-177A-24	Sequence 24, Appli
11	19	0.2	1325	4 US-08-913-795-6	Sequence 6, Appli
12	19	0.2	2040	1 US-08-599-252-103	Sequence 103, Appli
13	19	0.2	2040	5 PCT-US96-06352-103	Sequence 103, Appli
14	19	0.2	2040	5 PCT-US96-06583-103	Sequence 103, Appli
C 15	19	0.2	2169	3 US-08-806-326-5	Sequence 5, Appli
C 16	19	0.2	3286	4 US-09-211-417-2	Sequence 2, Appli
C 17	19	0.2	8797	2 US-08-723-306-6	Sequence 6, Appli
C 18	19	0.2	8797	5 PCT-US96-10041-6	Sequence 6, Appli
C 19	19	0.2	11093	2 US-08-723-306-5	Sequence 5, Appli
C 20	19	0.2	11093	5 PCT-US96-10041-5	Sequence 5, Appli
21	19	0.2	17056	4 US-09-245-041-3	Sequence 3, Appli
C 22	18	0.2	20	4 US-09-593-589-35	Sequence 35, Appli
23	18	0.2	47	1 US-08-222-177A-122	Sequence 122, Appli
24	18	0.2	47	1 US-08-222-177A-409	Sequence 409, Appli
25	18	0.2	50	1 US-08-222-177A-328	Sequence 328, Appli
26	18	0.2	51	1 US-08-222-177A-424	Sequence 424, Appli
27	18	0.2	65	1 US-08-222-177A-116	Sequence 116, Appli

28	18	0.2	65	1 US-08-222-177A-421	Sequence 421, App
29	18	0.2	72	1 US-08-222-177A-131	Sequence 131, App
30	18	0.2	72	1 US-08-222-177A-427	Sequence 427, App
31	18	0.2	88	3 US-08-545-860D-93	Sequence 93, Appl
32	18	0.2	109	3 US-08-532-896-29	Sequence 29, Appl
33	18	0.2	186	1 US-08-222-177A-12	Sequence 12, Appl
34	18	0.2	194	1 US-08-222-177A-15	Sequence 15, Appl
35	18	0.2	225	3 US-08-545-860D-91	Sequence 91, Appl
36	18	0.2	228	5 PCT-US94-04496-56	Sequence 56, Appl
37	18	0.2	228	5 PCT-US94-04496-56	Sequence 56, Appl
38	18	0.2	240	1 US-08-222-177A-30	Sequence 30, Appl
39	18	0.2	264	1 US-08-222-177A-10	Sequence 10, Appl
40	18	0.2	288	4 US-09-556-868-1	Sequence 1, Appli
41	18	0.2	298	1 US-08-599-252-88	Sequence 88, Appl
42	18	0.2	298	5 PCT-US96-06352-88	Sequence 88, Appl
43	18	0.2	298	5 PCT-US96-06583-88	Sequence 88, Appl
44	18	0.2	354	1 US-07-953-230A-7	Sequence 7, Appli
45	18	0.2	439	1 US-08-320-559-34	Sequence 34, Appli

ALIGNMENTS

RESULT 1
US-09-328-111-458
; Sequence 458, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertli, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 458
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(561)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-458

Query Match 4.0%; Score 389; DB 4; Length 561;
Best Local Similarity 100.0%; Pred. No. 3.5e-167;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	6691	accccaacaatttcaagccacagtcacagagaagtctcaggaaagcagcgtagaggaa	6750
Db	1	accccaacaatttcaagccacagtcacagagaagtctcaggaaagcagcgtagaggaa	60
QY	6751	gaatccttagcactcaggaagaacacacacatcagtaggaaagcagcgtatgacacacaaa	6810
Db	61	gaatccttagcactcaggaagaacacacacatcagtaggaaagcagcgtatgacacacaaa	120
QY	6811	ccagcaggagggtgatgagaagacatgaaagcatttatgggaactccagtcagaaattg	6870

NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06583
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,252
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 9053-0001.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US96-06583-105

Query Match 0.2%; Score 20; DB 5; Length 340;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4973 agactacacacacacaca 4992
|||||
Db 228 AGACTACACACACACACA 209

RESULT 5
US-08-991-789A-221/c
Sequence 221, Application US/08991789A
Patent No. 6225054
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 292
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,789A
FILING DATE: 11-Dec-1997
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 210121.419C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-991-789A-221

Query Match 0.2%; Score 20; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 377 gatctagctctctctgac 396
|||||
Db 182 GATCTAGCTCTCTCTGAC 163

RESULT 6
US-09-062-451-221/c
Sequence 221, Application US/09062451
Patent No. 6344550
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
APPLICANT: Smith, John M.
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 297
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,451
FILING DATE: 04-APR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.419C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-062-451-221

Query Match 0.2%; Score 20; DB 4; Length 530;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 gatctagctctctctgtac 396
|||||
Db 182 GATCTAGCTTCTCTCTGAC 163

RESULT 7
US-08-751-189-1
; Sequence 1, Application US/08751189
; Patent No. 5919656
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5919656el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,189
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7881 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-751-189-1

Query Match 0.2%; Score 20; DB 2; Length 7881;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4078 gaagaagcagtgctgtgg 4097
|||||
Db 4345 GAAGAAGCAGTGCTGTGG 4364

RESULT 8
US-09-060-836-1
; Sequence 1, Application US/09060836
; Patent No. 5981707
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5981707el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,836
; FILING DATE:

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7881 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-060-836-1

Query Match 0.2%; Score 20; DB 2; Length 7881;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4078 gaagaagcagtgctgtgg 4097
|||||
Db 4345 GAAGAAGCAGTGCTGTGG 4364

RESULT 9
US-09-184-445-1
; Sequence 1, Application US/09184445
; Patent No. 6174703
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 6174703el Genes Encoding Telomerase Protein
; TITLE OF INVENTION: 1
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,445
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7881 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna

```
US-09-184-445-1
Query Match          0.2%; Score 20; DB 4; Length 7881;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4078 gaagaagcagtggtgctgg 4097
      |||||||
Db 4345 GAAGAAGCAGTGGCTGG 4364

RESULT 10
US-08-222-177A-24
; Sequence 24, Application US/08222177A
; Patent No. 5582979
; GENERAL INFORMATION:
; APPLICANT: Weber, James L.
; TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
; TITLE OF INVENTION: (dC-dA)n.(dG-dT)n SEQUENCES AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 460
; CORRESPONDENCE ADDRES:
; ADDRESSEE: DeWitt Ross & Stevens, S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,177A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/341,562
; FILING DATE: 21-APR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: 09865.601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 831-2100
; TELEFAX: (608) 831-2106
; TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 259 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; INDIVIDUAL ISOLATE: Caucasian
; TISSUE TYPE: Blood
; IMMEDIATE SOURCE:
; CLONE: Mfd40
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 5
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 165..214
; OTHER INFORMATION: /rpt_type= "tandem"
; OTHER INFORMATION: /rpt_family= "(dC-dA)n.(dG-dT)n"
; OTHER INFORMATION: /citation= ([2])
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 65..85
; IDENTIFICATION METHOD: experimental
```

; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/42983
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: Mouse Lung
; US-08-915-795-6

Query Match 0.2%; Score 19; DB 4; Length 1325;

Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4978 acacacacacacagc 4996
|||||
DB 1274 ACACACACACAGAGC 1292

RESULT 12

; Sequence 103, Application US/08599252
; Patent No. 5705343
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIKKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,252
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US96-06352-103

Query Match 0.2%; Score 19; DB 5; Length 2040;

Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7887 cacacacacacaaagaa 7905
|||||
DB 887 CACACACACACAAGAA 905

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-599-252-103

Query Match 0.2%; Score 19; DB 1; Length 2040;

Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7887 cacacacacacacaaagaa 7905
|||||
DB 887 CACACACACACAAGAA 905

RESULT 13

; Sequence 103, Application PC/TUS9606352
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIKKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06352
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,252
; FILING DATE: 09-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US96-06352-103

Query Match 0.2%; Score 19; DB 5; Length 2040;

Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7887 cacacacacacacaaagaa 7905
|||||
DB 887 CACACACACACAAGAA 905

RESULT 14
PCT-US96-06583-103
; Sequence 103, Application PC/TUS9606583
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIKKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; TITLE OF INVENTION: HEMOCHROMATOSIS
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: PCT/US96/06583
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,252
; FILING DATE: 09-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US96-06583-103

Query Match 0.2%; Score 19; DB 5; Length 2040;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7887 cacacacacacacacaa 7905
|||||
DB 887 CACACACACACACAAAGAA 905

RESULT 15
US-08-806-326-5/c
; Sequence 5, Application US/08806326
; Patent No. 6022738
; GENERAL INFORMATION:
; APPLICANT: Atweb, George F.
; TITLE OF INVENTION: VECTORS FOR GENE THERAPY OF ERYTHROID
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA

; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/806,326
; FILING DATE: 26-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/398,160
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Richard S.
; REGISTRATION NUMBER: 26,154
; REFERENCE/DOCKET NUMBER: A30017-165/30389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2558
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2169 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2169
; US-08-806-326-5

Query Match 0.2%; Score 19; DB 3; Length 2169;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4978 acacacacacacacagc 4996
|||||
DB 1484 ACACACACACACACAGC 1466

Search completed: August 24, 2002, 01:12:05
Job time: 26825 sec

THIS PAGE BLANK (USPTO)